

Thouless-Anderson-Palmer equations for conditional Gibbs measures in the generic p -spin glass model

Antonio Auffinger ^{*}
Northwestern University

Aukosh Jagannath [†]
University of Toronto

December 20, 2016

Abstract

We study the Thouless-Anderson-Palmer (TAP) equations for spin glasses on the hypercube. First, using a random, approximately ultrametric decomposition of the hypercube, we decompose the Gibbs measure, $\langle \cdot \rangle_N$, into a mixture of conditional laws, $\langle \cdot \rangle_{\alpha, N}$. We show that the TAP equations hold for the spin at any site with respect to $\langle \cdot \rangle_{\alpha, N}$ simultaneously for all α . This result holds for generic models provided that the Parisi measure of the model has a jump at the top of its support. Second, we derive the TAP equations for spin distributions. We obtain an alternative representation for spin distributions that satisfy the cavity equations. This representation provides expressions for spin statistics as solutions of partial differential equations.

1 Introduction

The Thouless-Anderson-Palmer (TAP) equations were introduced by Thouless, Anderson, and Palmer [23] as the mean field equations for the Sherrington-Kirkpatrick (SK) model of spin glasses. These equations can be stated informally as follows. For each $\sigma \in \Sigma_N = \{-1, 1\}^N$, let

$$H_N(\sigma) = \frac{1}{\sqrt{N}} \sum_{i,j=1}^N g_{ij} \sigma_i \sigma_j$$

be the Hamiltonian for the SK model. Here g_{ij} are i.i.d. standard Gaussian random variables. Let

$$\mu_N(\{\sigma\}) = \frac{e^{-\beta H_N(\sigma) + h \sum_{i=1}^N \sigma_i}}{Z_N}$$

be the Gibbs measure of this system at inverse temperature, β , and external field, h . Here β and h are non-negative real numbers and Z_N is chosen such that μ_N is a probability measure on Σ_N . We denote integration of a quantity, say σ_i , against μ_N as $\langle \sigma_i \rangle$. The TAP equations state that in the limit that $N \rightarrow \infty$, we have that

$$\langle \sigma_i \rangle_\alpha \approx \tanh \left(h + \left\langle \frac{1}{\sqrt{N}} \beta \sum_j g_{ij} \sigma_j \right\rangle_\alpha - \beta^2 (1 - q_*) \langle \sigma_i \rangle_\alpha \right), \quad (1.1)$$

^{*}Department of Mathematics, Northwestern University, auffing@math.northwestern.edu

[†]Department of Mathematics, University of Toronto, aukosh@math.toronto.edu

for some $q_* \in [0, 1]$ and for some random measure for which integration is denoted by $\langle \cdot \rangle_\alpha$.

There have been two approaches to proving the TAP equations rigorously. The first approach is to take $\langle \cdot \rangle_\alpha$ as integration with respect to the Gibbs measure. This has been done by Talagrand [21] and Chatterjee [8] at sufficiently high temperature for the SK model where they establish (1.1) under this interpretation. A second approach, introduced by Bolthausen [6], is to interpret $\langle \sigma_i \rangle_\alpha$ as a vector in high dimensions, and to understand (1.1) through a fixed point iteration scheme. There he showed that this iteration converges to a unique solution of (1.1) in the entire predicted high temperature regime. At low temperature, as far we know, there is no rigorous proof of (1.1). In this regime, it is expected that there are many distinct measures, $\mu_{\alpha,N}$, called “pure states”, whose convex combination is μ_N and each of which satisfies (1.1).

The first goal of this paper is to study (1.1) for generic mixed p -spin glasses without an assumption on the temperature. These models are defined as follows. Consider the mixed p -spin glass Hamiltonian, $H_N(\sigma)$, which is the centered Gaussian process on $\Sigma_N = \{-1, 1\}^N$ with covariance

$$\mathbb{E}H_N(\sigma^1)H_N(\sigma^2) = N\xi(R_{12}),$$

where $R_{12} = \frac{1}{N} \sum \sigma_i^1 \sigma_i^2$ is called the overlap and $\xi(t) = \sum_{p \geq 2} \beta_p t^p$ is called the model. We let μ_N denote the corresponding Gibbs measure and $\langle \cdot \rangle$ expectation under products of μ_N . The SK model corresponds to $\xi(t) = \beta_2 t^2$. A mixed p -spin glass model is called generic if the set $\{t^p : \beta_p > 0\}$ is total in $(C([0, 1]), \sup |\cdot|)$.

Denote by ζ_N the distribution of the overlap under the measure $\mathbb{E}\mu_N^{\otimes 2}$, that is,

$$\zeta_N(A) = \mathbb{E}\langle \mathbf{1}(R_{12} \in A) \rangle$$

for any measurable $A \subset [-1, 1]$. It is known that ζ_N converges to ζ , where ζ is the unique minimizer of the Parisi formula [4, 17]. It is also known that generic models satisfy the Ghirlanda-Guerra identities in the limit [15, 17]. As a result, their asymptotic Gibbs measures [2] are known to have ultrametric support by Panchenko’s ultrametricity theorem [16]. We assume that ζ has a jump at the top of its support. That is, if $q_* := \sup \text{supp}(\zeta)$, we assume that

$$\zeta(\{q_*\}) > 0. \tag{1.2}$$

This assumption is expected to hold in a wide range of models at all temperatures. For more on this see Remark 1.4.

This ultrametric structure is the starting point for our study of the analogue of (1.1) for generic models. It was shown in [10] that, as a consequence of Panchenko’s ultrametricity theorem, Σ_N can be decomposed as the disjoint union of a collection of clusters, $\{C_{\alpha,N}\}_{\alpha \in \mathbb{N}}$, which satisfy certain ultrametric-type properties. Heuristically, these clusters are essentially balls of radius q_* . Within a cluster, the points are at overlap roughly q_* , between clusters the points have overlap less than $q_* - o_N(1)$ with high μ_N probability. We recall the precise definition of these sets in Appendix A.1. A similar decomposition was obtained by Talagrand in [22].

For each of these clusters, $C_{\alpha,N}$, we define

$$\mu_{\alpha,N}(\cdot) := \mu_N(\cdot | C_{\alpha,N}). \tag{1.3}$$

That is, $\mu_{\alpha,N}$ is the Gibbs measure conditioned on the set $C_{\alpha,N}$, with the convention that if $C_{\alpha,N} = \emptyset$, then $\mu_{\alpha,N} = \delta_{(1,\dots,1)}$. This yields a decomposition of the Gibbs measure μ_N as

$$\mu_N(\cdot) = \sum_{\alpha} \mu_{\alpha,N}(\cdot) \mu_N(C_{\alpha,N}) + o_N(1). \tag{1.4}$$

Integration with respect to the conditional measure $\mu_{\alpha,N}$ will be denoted by $\langle \cdot \rangle_{\alpha,N}$.

We now state our main theorem, which is the equivalent of (1.1) for generic models. For $\sigma \in \Sigma_N$, let

$$s(\sigma) = \sigma_1, \quad y_N(\sigma) = \sum_{p \geq 2} \frac{\beta_p}{N^{\frac{p-1}{2}}} \sum_{i_2, \dots, i_p} J_{i_2 \dots i_p} \sigma_{i_2} \cdots \sigma_{i_p}, \quad (1.5)$$

with $J_{i_2 \dots i_p} = g_{1i_2 \dots i_p} + g_{i_2 1 \dots i_p} + \dots + g_{i_2 \dots i_p 1}$. We call $s(\sigma)$ the spin of the first particle and y_N the local field on the first particle. Note that y_N is a centered Gaussian process on Σ_{N-1} with covariance given by

$$\mathbb{E} y_N(\sigma^1) y_N(\sigma^2) = \xi'(N^{-1}(\sigma^1, \sigma^2)). \quad (1.6)$$

For more on y_N see Lemma A.6. We also note here that the choice of the first spin as opposed to any fixed i will be irrelevant by site symmetry.

Our first result is that the TAP equation for a spin holds for the measures $\langle \cdot \rangle_{\alpha,N}$.

Theorem 1.1. *Assume that $\zeta(q_*) > 0$. We have that*

$$\left(\langle s \rangle_{\alpha,N} - \tanh \left[\langle y_N \rangle_{\alpha,N} + h - (\xi'(1) - \xi'(q_*)) \langle s \rangle_{\alpha,N} \right] \right)_{\alpha \in \mathbb{N}} \rightarrow 0 \quad (1.7)$$

in distribution.

The proof of Theorem 1.1 has several steps and along the way we pick up results that are of independent interest. We will outline the proof of Theorem 1.1 in the next section. We conclude this section with the following remarks regarding Theorem 1.1.

Remark 1.2. At high temperature, when the Parisi measure $\zeta = \delta_0$, the decomposition $C_{\alpha,N}$ is given by $C_{1,N} = \Sigma_N$, $C_{\alpha,N} = \emptyset$, $\alpha > 1$. The conditional measure $\mu_{1,N}$ is now identical to the Gibbs measure μ_N and one recovers the result of Talagrand [21] for a single spin.

Remark 1.3. Theorem 1.1 establishes the TAP equations for a single spin. The TAP equations are also predicted to hold for all spins $\sigma_1, \dots, \sigma_N$ simultaneously.

Remark 1.4. The assumption that the Parisi measure has a jump at the top of its support, $\zeta(q_*) > 0$, is natural and believed to be true for a large collection of (if not all) generic models at all temperatures. Results in this direction were obtained by Auffinger-Chen (see Theorem 4 in [3]). If there is no jump at the top of the support, then it is unclear the extent to which a true pure state decomposition will hold in such systems [19]. As we shall see shortly, however, at infinite particle number, (1.1) can be shown without this assumption.

Remark 1.5. Since the statement of Theorem 1.1 depends on the construction of the measures $\langle \cdot \rangle_{\alpha,N}$, one may wonder what would happen if one takes a different decomposition. In Section 5 we show that the decomposition (1.3) is essentially unique in the following sense. Any other collection of subsets $X_{\alpha,N}$ that satisfy the same properties as $C_{\alpha,N}$ must also satisfy $\mu_N(X_{\alpha,N} \Delta C_{\alpha,N}) \rightarrow 0$.

1.1 Outline of the proof of Theorem 1.1

Theorem 1.1 relates the quantities

$$\langle s \rangle_{\alpha,N} = \frac{1}{\mu C_{\alpha,N}} \int_{C_{\alpha,N}} \sigma_1 d\mu_N \quad \text{and} \quad \langle y_N \rangle_{\alpha,N} = \frac{1}{\mu C_{\alpha,N}} \int_{C_{\alpha,N}} y_N(\sigma) d\mu_N.$$

Put differently, we are interested in the relation between s and y within a cluster, C_α . Heuristically, for large N there is little difference between a fixed coordinate and a “cavity coordinate”. By a cavity coordinate, we mean that we study the law of $(s_{\alpha,N}, y_{\alpha,N})$ which are distributed like $(\epsilon, y_N(\sigma))$ drawn from a tilted measure on Σ_{N+1} ,

$$\mu_N^\top = \frac{e^{\epsilon y_N(\sigma)} d\mu_N(\sigma)}{\int 2 \cosh(y_N(\sigma)) d\mu_N}$$

conditioned on the event $\{\sigma \in C_{\alpha,N}\}$. Call this conditional measure $\mu_{\alpha,N}^\top$. In this step, we will assume that y_N is independent of μ_N and satisfies

$$\mathbb{E} y_N(\sigma^1) y_N(\sigma^2) = \xi'(R_{12}) + o_N(1).$$

As a result, to study convergence of $(s_{\alpha,N}, y_{\alpha,N})$ for a fixed α , it suffices to study convergence of statistics of the form

$$\mathbb{E} \prod_i \int \phi_i(\epsilon, y_N) d\mu_{\alpha,N}^\top$$

for any finite family of reasonable ϕ_i . These statistics, as we will find, are continuous functionals of the law of the overlap array of i.i.d. draws from $\mu_{\alpha,N}$. The $\mu_{\alpha,N}$ are asymptotically *replica symmetric*, that is, their overlap array converges to the matrix which is 1 on the diagonal and $q_* = \sup \text{supp}\{\zeta\}$ on the off diagonal. This implies that the law of $(s_{\alpha,N}, y_{\alpha,N})$ converges to the law of a stochastic process (s, y) at infinite particle number, which can be described as follows: let h_α be a centered gaussian with variance $\xi'(q_*)$. Then (s, y) are the random variables with conditional density

$$p(s, y; h_\alpha) \propto e^{-\frac{(y-h_\alpha)^2}{2(\xi'(1)-\xi'(q_*))}} e^{sy}, \quad (1.8)$$

with respect to the product of the counting measure on Σ_1 and Lebesgue measure on \mathbb{R} . It is an elementary calculation to show that this satisfies the TAP equations, conditionally on h_α . Indeed, once making this reduction, this is similar in spirit to the high temperature setting as in [8]. This is stated and proved in a slightly more general setting in [8] and also Theorem 1.9. This is shown in Section 3.

The question is then: “to what extent can we treat a fixed coordinate as a cavity coordinate”. To understand the main issue, fix a coordinate, say $i = 1$. (By site symmetry this is without loss of generality.) A standard argument (Lemma A.6) shows that for $\rho = (\epsilon, \sigma) \in \Sigma_1 \times \Sigma_{N-1}$, one has the approximation

$$H_N(\rho) \approx H'_N(\sigma) + \epsilon y_N(\sigma)$$

for some Hamiltonian H'_N on Σ_{N-1} . There are three obstacles to this that were not present in the setting of a cavity coordinate:

1. The sets $C_{\alpha,N}$ are not products.
2. The variable $y_N(\sigma)$ is correlated to μ_N .
3. The variable $y_N(\sigma)$ is correlated to $C_{\alpha,N}$.

We begin by tackling the first two obstacles. To this end, we construct $\{\tilde{W}_{\alpha,N}\}$, subsets of Σ_{N-1} , using Theorem A.1 applied to μ'_N , the Gibbs measure for H' . We lift these sets to obtain $W_{\alpha,N}^\dagger =$

$\tilde{W}_{\alpha,N} \times \Sigma_1 \subset \Sigma_N$. We then permute these to be in decreasing μ_N mass. Call this permutation map π_N , and the permuted sets $W_{\alpha,N} = W_{\pi_N(\alpha),N}^\dagger$. In Section 4, we show that these sets satisfy the conditions of Theorem A.1. We then find that an essential uniqueness holds for such sets which implies that μ_N conditioned on $C_{\alpha,N}$ is essentially μ_N conditioned on $W_{\alpha,N}$. This is shown in Section 5.

The problem then reduces to controlling quantities like

$$\int_{W_{\alpha,N}} y_N d\mu_N \approx \int_{\tilde{W}_{\pi(\alpha),N}} y_N d(\mu'_N)^\top,$$

where we use the notation of the cavity coordinate from above. Evidently, the first two obstructions are now avoided. Furthermore, the correlation between $\{C_{\alpha,N}\}$ and y_N is contained in the map π_N . By the discussion above, (s, y) is like that for a cavity coordinate with respect to $(\mu'_N)^\top$. The terms in the display above are simply re-arrangements (through π_N) of the terms we wish to study. Thus, since this collection converges to zero, Theorem 1.1 will hold provided π_N is not too wild. This is shown in Section 6.

1.2 Results at Infinite Particle number

We now turn to our results at infinite particle number. The main objects of study in this section are what are called *spin distributions*, introduced by Panchenko in [18], which are a natural family of limit objects for spins of draws from a Gibbs measure. Let \mathcal{M} be the set of all measures on the set $\{-1, 1\}^{\mathbb{N} \times \mathbb{N}}$ that are exchangeable, that is, if (s_i^ℓ) has law $\nu \in \mathcal{M}$, then

$$(s_{\pi(i)}^{\rho(\ell)}) \stackrel{(d)}{=} (s_i^\ell)$$

for any permutations π, ρ of the natural numbers. The Aldous-Hoover theorem [17, Theorem 1.4], states that if (s_i^ℓ) is the random variable induced by some measure $\nu \in \mathcal{M}$, then there is a measurable function of four variables, $\sigma(w, u, v, x)$, such that

$$(s_i^\ell) \stackrel{(d)}{=} (\sigma(w, u_\ell, v_i, x_{\ell i}))$$

where $w, u_\ell, v_i, x_{\ell i}$ are i.i.d. uniform $[0, 1]$ random variables. We call this function a *directing function* for ν .

For any ν in \mathcal{M} with directing function σ , let $\bar{\sigma}(w, u, v) = \int \sigma(w, u, v, x) dx$. Note that since σ is $\{\pm 1\}$ -valued, this encodes all of the information of $\sigma(w, u, v, \cdot)$. Define the measure μ on the Hilbert space, $\mathcal{H} = L^2([0, 1], dv)$, by the push-forward of du through the map $u \mapsto \bar{\sigma}(w, u, \cdot)$,

$$\mu = (u \mapsto \bar{\sigma}(w, u, \cdot))_* du.$$

We call μ the asymptotic Gibbs measure corresponding to ν . Finally, consider *the overlap array*, which is the array

$$R_{\ell\ell'} = (\bar{\sigma}(w, u_\ell, \cdot), \bar{\sigma}(w, u_{\ell'}, \cdot))_{\mathcal{H}}. \quad (1.9)$$

Heuristically, one should think of $(\sigma(w, u_\ell, \cdot, \cdot))$ as limits of configuration vectors that are independently drawn from the Gibbs measure of the system, with $(R_{\ell\ell'})$ being the limit of the overlap array of these draws, s_i^ℓ being the limit of spin at site i for the ℓ -th draw, and $\bar{\sigma}(w, u_\ell, v_i)$ being the

magnetization at this site. For a more precise explanation of spin distributions see [17, 18], and for the notion of asymptotic Gibbs measures see [2] and [17].

We are interested in measures, ν , that satisfy two common invariance properties: the Ghirlanda-Guerra identities and the cavity equations. (See Appendix A.5 for the definition of these properties.) As an example, systems arising from Hamiltonians with regularizing perturbations [17] satisfy this property, as well as the generic p -spin models (see Appendix A.6). We let \mathcal{M}_{inv}^ξ denote the space of ν that satisfy the cavity equations and the Ghirlanda-Guerra identities. The goal of this section is to obtain an alternative representation formula for such measures, which provides expressions for moments of products of spins (s_i^ℓ) as the solutions of certain partial differential equations.

We begin by defining the state measure, which describes the law of the spin within a configuration.

Definition 1.6. (State Measure) For any $\sigma \in \mathcal{H}$ with $\|\sigma\|_{\mathcal{H}} = \sqrt{q_*}$ and any measurable $f(\sigma)$, the *state at σ with boundary data f* to be the probability measure, μ_σ^f , on $\{-1, 1\} \times \mathbb{R}$ which is defined as follows. For every continuous bounded function, $\phi = \phi(s, y)$, we have

$$\int \phi d\mu_\sigma^f := \int_{\mathbb{R}} \sum_s \phi(s, y) p(s, y; f(\sigma)) dy \quad (1.10)$$

where $p(s, y; h)$ is defined as in (1.8).

We denote the expectation above as $\langle \phi \rangle_\sigma^f$, which we call the *state average of ϕ* . When it is unambiguous we omit the superscript for the boundary data. For multiple copies, $(s_i, y_i)_{i=1}^\infty$, drawn from the product $\mu_\sigma^{\otimes \infty}$, we also denote the average by $\langle \cdot \rangle_\sigma$.

Informally, this will allow for a representation of spins as follows: for a random measure μ on \mathcal{H} , one draws a configuration σ and then forms the spins for this configuration as i.i.d draws from μ_σ^f for some suitable choice of data f . For the reader familiar with Gibbs measures in the sense of Dobrushin-Lanford-Ruelle, think of this as the conditional law of a site where the function f plays the role of the Hamiltonian given the data for the complement of this site.

We will be most interested in the setting where f is equal to one of two specific choices of data which we define presently. Fix ν in \mathcal{M}_{inv}^ξ . Let μ be the asymptotic Gibbs measure for ν , and let σ be its directing function. Let ζ be the law of R_{12} with respect to $\mathbb{E}\mu^{\otimes 2}$. Let q_* be the supremum of the support of ζ . Let U be an ultrametric subset of the sphere of radius $\sqrt{q_*}$ in \mathcal{H} that is positive in the sense that for any two $x, y \in U$, we have $\langle x, y \rangle \geq 0$. A consequence of Panchenko's ultrametricity theorem, combined with other well-known properties of the Ghirlanda-Guerra identities, is that the support of μ is such a set [17].

Define the driving process on U to be the Gaussian process, $B_t(\sigma)$, indexed by $(t, \sigma) \in [0, q_*] \times U$, which is centered, a.s. continuous in time and measurable in space, with covariance

$$\text{Cov}_B((t_1, \sigma^1), (t_2, \sigma^2)) = (t_1 \wedge t_2) \wedge (\sigma^1, \sigma^2). \quad (1.11)$$

Put concretely, for each fixed σ , $B_t(\sigma)$ is a Brownian motion and for finitely many (σ^i) , $(B_t(\sigma^i))$ is a family of branching Brownian motions whose branching times are given by the inner products between these σ^i .

We then define the cavity field process on U as the solution, $Y_t(\sigma)$, of the SDE

$$\begin{cases} dY_t(\sigma) = \sqrt{\xi''(t)} dB_t(\sigma) \\ Y_0(\sigma) = h. \end{cases} \quad (1.12)$$

One should think of $Y_{q*}(\sigma)$ as the limit of y_N from (1.6). Indeed observe that it has the same covariance. Evidently, after comparing (1.8) with the state measure (1.10) with boundary data $f = Y_{q*}$, this will be one of the data of interest.

To define the other data of interest, we introduce the following. Let u be the unique weak solution to the Parisi initial value problem,

$$\begin{cases} u_t + \frac{\xi''(t)}{2} (u_{xx} + \zeta([0, t])u_x^2) = 0 & (0, 1) \times \mathbb{R} \\ u(1, x) = \log \cosh(x). \end{cases} \quad (1.13)$$

For the definition of weak solution in this setting and basic properties of u see [12]. We now define the local field process, $X_t(\sigma)$, to be the solution to the SDE

$$\begin{cases} dX_t(\sigma) = \xi''(t)\zeta([0, t])u_x(t, X_t(\sigma))dt + dY_t(\sigma) \\ X_0(\sigma) = h. \end{cases} \quad (1.14)$$

Finally, let the magnetization process be $M_t(\sigma) = u_x(t, X_t(\sigma))$. We will show that the process $X_{q*}(\sigma)$ is related to a re-arrangement of $Y_{q*}(\sigma)$. If we view σ as a state, then $M_{q*}(\sigma)$ will be the magnetization of this state. The process $X_{q*}(\sigma)$ will be the other data of interest.

The basic properties of these processes, e.g., existence, measurability, continuity, etc, are studied briefly in Appendix A.2. We invite the reader to compare their definitions to [13, Eq. IV.51] and [7, Eq. 0.20]. Finally, we observe that by an application of Girsanov's theorem (see specifically [11, Lemma 8.3.1]), (1.10), can be equivalently written as

$$\int \phi d\mu_\sigma^f := \mathbb{E} \left(\frac{\sum_{s \in \{\pm 1\}} \phi(s, X_1) e_1^{X_1 s}}{2 \cosh(X_1)} \middle| X_{q*} = f(\sigma) \right). \quad (1.15)$$

Let μ be a random measure on \mathcal{H} such that the corresponding overlap array (1.9) satisfies the Ghirlanda-Guerra identities. Consider the law of the random variables (S, Y) defined through the relation:

$$\mathbb{E} \langle \phi(S, Y) \rangle = \mathbb{E} \int \langle \phi \rangle_\sigma^X d\mu \quad (1.16)$$

and the random variables (S', Y') defined through the relation

$$\mathbb{E} \langle \phi(S', Y') \rangle = \mathbb{E} \int \langle \phi \rangle_\sigma^Y \frac{\cosh(Y_{q*}(\sigma))}{\int \cosh(Y_{q*}(\sigma)) d\mu} d\mu.$$

Let (S_i, Y_i) be drawn from $(\mu_\sigma^X)^{\otimes \infty}$ and (S'_i, Y'_i) be drawn from $(\mu_\sigma^Y)^{\otimes \infty}$ where σ is drawn from μ . For i.i.d. draws $(\sigma^l)_{l \geq 1}$ from $\mu^{\otimes \infty}$, we define (S_i^l, Y_i^l) and $(S_i'^l, Y_i'^l)$ analogously.

The main result in this section is the following representation theorem for spins from cavity invariant measures.

Theorem 1.7. *We have the following.*

1. For any ξ and any asymptotic Gibbs measure μ , let $(\sigma_l)_{l \geq 1}$ be i.i.d. draws from μ , let (S_i^l, Y_i^l) and $(S_i'^l, Y_i'^l)$ be defined as above with $\sigma = \sigma^l$. Then these random variables are equal in distribution.

2. For any measure ν in \mathcal{M}_{inv}^ξ , let (s_i^ℓ) denote the array of spins and μ denote its corresponding asymptotic Gibbs measure. Let (S_i^l) be defined as above with $\zeta = \mathbb{E}\mu^{\otimes 2}((\sigma^1, \sigma^2) \in \cdot)$. Then we have

$$(s_i^l) \stackrel{(d)}{=} (S_i^l).$$

Remark 1.8. In [17, 18], Panchenko obtained a description of the laws of (s_i^l) in a finite replica symmetry breaking regime (i.e., ζ consists of finitely many atoms) using Ruelle probability cascades (see (2.1)). This is a key step in our proof of Theorem 1.7. At finite replica symmetry breaking, the connection to the process X_t can already be seen in [7, pp 249-250] as a consequence of the Bolthausen-Sznitman invariance principle.

As a consequence of this representation formula, we are able to obtain formulas for evaluations of functions of spins in terms of integrals against ζ and the solutions of certain PDEs. This is explained in Section 7. A second consequence of these equations is a means to interpret (1.1) at infinite particle number as in the following theorem.

Theorem 1.9. *For any σ in \mathcal{H} and any measurable initial data f , the state at σ with boundary data f satisfies*

$$\langle s \rangle_\sigma^f = \tanh(\langle y \rangle_\sigma^f - (\xi'(1) - \xi'(q)) \langle s \rangle_\sigma^f).$$

We end this section with the following remarks.

Remark 1.10. This is to be compared with the discussion in [13, Chapter III] by interpreting ‘pure state’ as ‘state measure’. Indeed, these objects satisfy the relations of the averages “ $\langle \cdot \rangle_\alpha$ ” as defined in [13, Chap. 3]. For example, observe that we have the independence property

$$\langle s_1 \dots s_n \rangle_\sigma = \prod_{i=1}^n \langle s_i \rangle_\sigma = \prod_{i=1}^n u_x(q_*, X_{q_*}^i(\sigma)).$$

Remark 1.11. In the infinite particle number setting, a notion of pure states has been studied by Panchenko in [19]. Under the assumption that μ is purely atomic, he calls σ in the support of μ a pure state. For general ζ , he defines a notion of approximate pure state by grouping σ ’s in a natural fashion. The state measures defined above are kernels whose input are the elements of the support of μ , rather than being the elements themselves.

Remark 1.12. Note that the dependence on h in Theorem 1.9 is implicitly by suitable choice of f .

Acknowledgements

The authors thank Dmitry Panchenko for numerous comments, suggestions and several discussions on a first version of this project, which dramatically improved the results of this paper. We also thank Gérard Ben Arous and Ian Tobasco for several fruitful discussions. A.A. would also thank Louis-Pierre Arguin, Wei-Kuo Chen and Nicola Kistler for helpful and broad discussions about TAP. A.J. thanks the Northwestern University for their hospitality. This research was conducted while A.A. was supported by NSF DMS-1597864 and A.J. was supported by NSF OISE-1604232.

Notation

For a vertex α of a tree, we denote by $|\alpha|$ the depth of α , that is its (edge or vertex) distance from the root. We denote by $p(\alpha)$ to be the set of vertices in the path from the root to α . For two vertices α, β , we let $\alpha \wedge \beta$ denote their least common ancestor, and we say that $\alpha \preceq \beta$ if $\alpha \in p(\beta)$. In particular $\alpha \preceq \alpha$. We say that $\beta \in \text{child}(\alpha)$ if $|\beta| = |\alpha| + 1$ and $\alpha \preceq \beta$. We say that $\alpha \approx \beta$ if neither $\alpha \preceq \beta$ nor $\beta \preceq \alpha$.

For any Polish space E we let $\text{Pr}(E)$ denote the set of Borel probability measures on this space, which we always equip with the topology of weak convergence of measures. Let \mathcal{Q}_d denote the set of $d \times d$ matrices of the form

$$\mathcal{Q}_d = \{(q_{ij})_{i,j \in [d]} : q_{ij} \in [0, 1], q_{ij} = q_{jk}, q_{ij} \geq q_{ik} \wedge q_{kj} \forall i, j, k\}.$$

Note that this set is a compact subset of \mathbb{R}^{d^2} in the induced topology. We say that $f \lesssim_a g$ if there exists a positive constant $C(a)$ that depends at most on a such that $f(x) \leq C(a)g(x)$ for all x in the intersection of the domains of f and g . If E is a measurable set and f an integrable function then $\int_E f d\mu = \frac{1}{\mu(E)} \int f d\mu$ with the convention that this is zero if $\mu(E) = 0$.

2 Proofs at Infinite Particle Number

We now turn to the proofs of the results at infinite particle number. Before we can state these results we need to recall certain basic results of Panchenko from the theory of spin distributions [17, 18].

2.1 Preliminaries

We begin with the observation that if we apply the cavity equations, (A.6) with $n = m$ and $r = 0$, we get that,

$$\mathbb{E} \prod_{l \leq q} \prod_{i \in C_l} s_i^l = \mathbb{E} \frac{\prod_{l \leq q} \mathbb{E}' \prod_{i \in C_l} \tanh(G_{\xi', i}(\bar{\sigma})) \prod_{i \leq n} \cosh(G_{\xi', i}(\bar{\sigma}))}{\left(\mathbb{E}' \prod_{i \leq n} \cosh(G_{\xi', i}(\bar{\sigma})) \right)^q}.$$

Note that the righthand side is a function of only the overlap distribution of $\bar{\sigma}$ corresponding to ν . Let the law of R_{12} be denoted by ζ . Suppose that ζ consists of $r + 1$ atoms. Then, since μ satisfies the Ghirlanda-Guerra identities by assumption, we know that this can also be written as

$$\mathbb{E} \prod_{l \leq q} \prod_{i \in C_l} s_i^l = \mathbb{E} \frac{\prod_{l \leq q} \sum_{\alpha} w_{\alpha} \prod_{i \in C_l} \tanh(g_{\xi', i}(h_{\alpha})) \prod_{i \leq n} \cosh(g_{\xi', i}(\bar{\sigma}))}{\left(\sum w_{\alpha} \prod_{i \leq n} \cosh(g_{\xi', i}(\bar{\sigma})) \right)^q}. \quad (2.1)$$

Here, $(w_{\alpha})_{\alpha \in \partial \mathcal{A}_r}$ are the weights corresponding to a $RPC(\zeta)$ and $\{h_{\alpha}\}_{\alpha \in \mathcal{A}_r}$ are the corresponding vectors, with $\mathcal{A}_r = \mathbb{N}^r$ viewed as a tree with r levels.

In this setting, it is well known that $g_{\xi'}(h_{\alpha})$ has the following explicit version. Let $(\eta_{\alpha})_{\alpha \in \mathcal{A}_r}$ be i.i.d. gaussians, then

$$g_{\xi'}(h_{\alpha}) = \sum_{\beta \preceq \alpha} \eta_{\beta} (\xi'(q_{|\beta|}) - \xi'(q_{|\beta|-1}))^{1/2}.$$

It was then showed by Panchenko that the above also has the following representation in terms of “tilted” variables η' as follows.

We define the following family of functions $Z_p : \mathbb{R}^p \rightarrow \mathbb{R}$ with $0 \leq p \leq r$ recursively as follows. Let

$$Z_r(x) = \log \cosh \left(\sum_{i=1}^r x_i (\xi'(q_i) - \xi'(q_{i-1}))^{1/2} \right)$$

and let

$$Z_p(x) = \frac{1}{\zeta([0, q_p])} \log \int \exp(\zeta([0, q_p]) \cdot Z_{p+1}(x, z)) d\gamma(z) \quad (2.2)$$

where $d\gamma$ is the standard gaussian measure on \mathbb{R} . We then define the transition kernels

$$K_p(x, dx_{p+1}) = \exp(\zeta([0, q_p]) (Z_{p+1}(x, x_{p+1}) - Z_p(x, x_{p+1}))) d\gamma(x_{p+1}).$$

Finally, we define η'_α as as the random variable with law $K_{|\alpha|}((\eta_\beta)_{\beta \prec \alpha}, \cdot)$. Finally define

$$g'_{\xi'}(h_\alpha) = \sum_{\beta \prec \alpha} \eta'_\beta (\xi'(q_{|\beta|}) - \xi'(q_{|\beta|-1}))^{1/2}.$$

Define $g'_{\xi', i}$ analogously. We then have the following proposition.

Proposition 2.1 (Panchenko [18]). *Let w_α be as above and let*

$$w'_\alpha = \frac{w_\alpha \prod_{i \leq n} \cosh(g'_{\xi', i}(h_\alpha))}{\sum w_\alpha \prod_{i \leq n} \cosh(g'_{\xi', i}(h_\alpha))}$$

Then we have

$$((w'_\alpha, g'_{\xi', i}(h_\alpha)))_\alpha \stackrel{(d)}{=} ((w_\alpha, g'_{\xi', i}(h_\alpha)))_\alpha.$$

If we apply this proposition to (2.1), we have that

$$\mathbb{E} \prod_{l \leq q} \prod_{i \in C_l} s_i^l = \mathbb{E} \prod_{l \leq q} \sum_{\alpha} w_\alpha \prod_{i \in C_l} \tanh(g'_{\xi', i}(h_\alpha)).$$

2.2 Proof of Theorem 1.7.

We now turn to proving Theorem 1.7. We begin with the following two lemmas.

Lemma 2.2. *Let h_α , η_α , g_ξ , g'_ξ be as above. We then have the following equalities in distribution*

$$\begin{aligned} (g(h_\alpha))_\alpha &\stackrel{(d)}{=} (B_{q_*}(h_\alpha))_\alpha \\ (g_{\xi'}(h_\alpha))_\alpha &\stackrel{(d)}{=} (Y_{q_*}(h_\alpha))_\alpha \\ (g'_{\xi'}(h_\alpha))_\alpha &\stackrel{(d)}{=} (X_{q_*}(h_\alpha))_\alpha. \end{aligned}$$

Proof. Observe by the independent increments property of Brownian motion, we have that

$$(\eta_\alpha) \stackrel{(d)}{=} (B(q_{|\alpha|}, h_\alpha) - B(q_{|\alpha|-1}, h_\alpha)).$$

This yields the first two equalities. It remains to see the last equality.

To this end, fix h_α , and consider the process X_t that solves the SDE (1.14). Then if Y_t is distributed like Y as above with respect to some measure Q , then by Girsanov's theorem [11, Lemma 8.3.1], we have that with respect to the measure P with Radon-Nikodym derivative

$$\frac{dP}{dQ}(t) = e^{\int_0^t \zeta([0,s]) du_s},$$

the process Y_t has the same law as X_t . In particular, for the finite collection of times q_i we have that

$$\begin{aligned} \mathbb{E}_P F(X_{q_0}, \dots, X_{q_r}) &= \int F(Y_{q_0}, \dots, Y_{q_r}) e^{\int_0^t \zeta du(s, Y_s)} dQ(Y) \\ &= \int F(Y_{t_1}, \dots, Y_{t_k}) \prod_{i=0}^k e^{\zeta([0, q_k])(u(q_k, Y_{q_k}) - u(q_{k-1}, Y_{q_{k-1}}))} dQ. \end{aligned}$$

By recognizing the law of (B_{q_k}) and (Y_{q_k}) as Gaussian random variables, and (2.2) as the Cole-Hopf solution of the Parisi IVP (1.13), $u(q_k, x) = Z_k(x)$, the result follows. \square

We now need the following continuity theorem. This is intimately related to continuity results commonly used in the literature, though the method of proof is different.

Let \mathcal{Q}_d denote the set of $d \times d$ matrices of the form

$$\mathcal{Q}_d = \{(q_{ij})_{i,j \in [d]} : q_{ij} \in [0, 1], q_{ij} = q_{jk}, q_{ij} \geq q_{ik} \wedge q_{kj} \forall i, j, k\}.$$

Note that this set is a compact subset of \mathbb{R}^{d^2} . Consider the space $\text{Pr}([0, 1])$ equipped with the weak-* topology. Then the product space $\text{Pr}([0, 1]) \times \mathcal{Q}_d$ is compact Polish. For any $Q \in \mathcal{Q}_d$, let $(\sigma^i(Q))_{i=1}^d \subset \mathcal{H}$ be a collection of vectors whose gram-matrix is Q . We can then define the functional

$$\mathcal{R}(\zeta, Q) = \mathbb{E} \prod_{i=1}^d u_x(q_*, X_{q_*}(\sigma^i)).$$

Lemma 2.3. *We have that \mathcal{R} is well-defined and is jointly continuous.*

Proof. Let $(\sigma^i)_{i=1}^d$ be any collection with overlap matrix Q . Recall the infinitesimal generator, L^{lf} , of the collection $(X_t(\sigma^i))$ from (A.1). Observe that L^{lf} depends on (σ^i) only through their overlap matrix, which is Q . Thus the law is determined by this matrix and \mathcal{R} is well-defined.

We now turn to proving continuity. As $\text{Pr}([0, 1]) \times \mathcal{Q}$ is compact Polish, it suffices to show that for $\zeta_r \rightarrow \zeta$ and $Q^r = (q_{ij}^r)$ with $q_{ij}^r \rightarrow q_{ij}$, $1 \leq i, j \leq l$,

$$\mathcal{R}(\zeta_r, Q^r) \rightarrow \mathcal{R}(\zeta, Q),$$

as $r \rightarrow \infty$.

Let a_{ij}^r and b_i^r be the coefficients of the diffusion associated to the local field process X^{ζ_r, Q^r} . By (A.1), we have

$$a_{ij}^r(t) = \mathbf{1}_{\{t \leq q_{ij}^r\}}, \quad b_i^r(t, \cdot) = \xi'' \zeta_r u_x^r(t, \cdot),$$

where u^r is the solution to the Parisi initial value problem corresponding to ζ_r . These coefficients are all uniformly bounded, measurable in time and smooth in space. Furthermore, ξ is continuous, so that

$$\begin{aligned} & \int_0^t (|a_{ij}^r(s) - a_{ij}(s)| + \sup_x |b^r(s, x) - b(s, x)|) ds \\ & \leq |q_{ij}^r - q_{ij}| + \int_0^t \sup_x |\zeta_r([0, s]) u_x^r(t, x) - \zeta([0, s]) u_x(t, x)| ds \rightarrow 0 \quad (2.3) \end{aligned}$$

as $r \rightarrow \infty$ since u_x^r converges uniformly to u_x by [3, Prop. 1] as $\zeta_r \rightarrow \zeta$.

By Stroock-Varadhan's theorem [20, Theorem 11.1.4], the convergence from (2.3) implies that the laws of the solutions to the corresponding martingale problems converge. As $(x_1, \dots, x_d) \mapsto \prod_{i=1}^d \tanh(x_i)$ is a continuous bounded function we obtain the continuity of F . \square

We may now turn to the proof of the main theorem of this section.

Proof of Theorem 1.7. Suppose first that ζ consists of $r+1$ atoms. In this setting the result has already been proved by the aforementioned results of Panchenko combined with Lemma 2.2. The main task is to prove these results for general ζ . To this end, let $\zeta_r \rightarrow \zeta$ be atomic. Denote the spins corresponding to these measures by $s_{i,r}^l$.

Correspondingly, for any collection of moments we have

$$\mathbb{E} \prod_{l \leq q} \prod_{i \in C_l} s_{i,r}^l = \mathbb{E} \langle \mathcal{R}(\zeta_r, Q) \rangle.$$

Recall that the overlap distribution converges in law when $\zeta_r \rightarrow \zeta$, thus by Lemma 2.3 and a standard argument,

$$\mathbb{E} \langle \mathcal{R}(\zeta_r, Q) \rangle_r \rightarrow \mathbb{E} \langle \mathcal{R}(\zeta, Q) \rangle = \mathbb{E} \left\langle \prod_{i \leq q} \prod_{i \in C_l} \tanh(X_{q_*}^i(\sigma^l)) \right\rangle.$$

However, as the overlap distribution determines the spin distribution, we see that

$$\mathbb{E} \prod_{l \leq q} \prod_{i \in C_l} s_{i,r}^l \rightarrow \mathbb{E} \prod_{l \leq q} \prod_{i \in C_l} s_i^l = \mathbb{E} \frac{\left\langle \prod_{l \leq q} \prod_{i \in C_l} \tanh(Y_{q_*}^i(\sigma^l)) \prod_{i \leq n} \cosh(Y_{q_*}^i(\sigma^l)) \right\rangle}{\left\langle \prod_{i \leq n} \cosh(Y_{q_*}^i(\sigma^l)) \right\rangle^q}.$$

This yields both results. \square

2.3 Proof of Theorem 1.9

We now prove that the TAP equation holds at infinite particle number. Before stating this proof we point out two well-known [4, 11] but useful facts: the magnetization process $u_x(s, X_s(\sigma))$ is a martingale for fixed σ and $u_x(t, x) = \tanh(x)$ for $t \geq q_* = \text{supp}\zeta$.

Proof of Theorem 1.9. Observe first that by Definition 1.6

$$\langle s \rangle_\sigma = \tanh(X_{q_*}(\sigma)).$$

Now for the local field, we have that again by Definition 1.6 and (1.14),

$$\begin{aligned}\langle y \rangle_\sigma &= \langle X_1 \rangle_\sigma = \mathbb{E}(X_1 | \mathcal{F}_{q_*}) \\ &= X_{q_*} + \int_{q_*}^1 \zeta[0, t] \xi''(t) \mathbb{E}(u_x(s, X_s) | \mathcal{F}_{q_*}) dt \\ (Mart.) &= X_{q_*} + (\xi'(1) - \xi'(q_*)) u_x(q_*, X_{q_*}),\end{aligned}$$

where in the last equality we use the Martingale property of the process $M_t(\sigma) = u_x(t, X_t(\sigma))$. Since $\zeta([0, t]) = 1$ for all $t \in [q_*, 1]$, we obtain $u_x(q_*, X_{q_*}) = \tanh(X_{q_*})$ and

$$\langle y \rangle_\sigma = X_{q_*} + (\xi'(1) - \xi'(q_*)) \tanh(X_{q_*}) = X_{q_*} + (\xi'(1) - \xi'(q_*)) \langle s \rangle_\sigma.$$

Putting this together, we see that

$$\langle s \rangle_\sigma = \tanh(X_{q_*}) = \tanh(\langle y \rangle_\sigma - (\xi'(1) - \xi'(q_*)) \langle s \rangle_\sigma)$$

as desired. \square

3 Convergence of Spins and Local fields for a Cavity Coordinate

In this section, we study the joint law of a spin and the local field on that spin for a cavity coordinate. For this terminology see Section 1.1. As a consequence of this, we will find that (1.7) will hold for a cavity coordinate.

Note: In the remainder of this paper we take $h = 0$. This does not change the arguments, however it simplifies the notation dramatically.

Let $(H'(\sigma))$ be a centered Gaussian process on Σ_N with covariance

$$\mathbb{E} H'(\sigma^1) \cdot H'(\sigma^2) = N \xi(R_{12}) + o_N(1) \quad (3.1)$$

where the $o_N(1)$ term is a function of the overlap as well. Let ν_N denote the Gibbs measure on Σ_N corresponding to H' . Let $(y(\sigma))$ be a centered Gaussian process on Σ_N that is independent of H' and satisfies

$$\begin{cases} \mathbb{E} y(\sigma^1) y(\sigma^2) = \xi'(R_{12}) + o_N(1) \\ y \perp\!\!\!\perp H', \end{cases} \quad (3.2)$$

where again the $o_N(1)$ term is a function of the overlap.

Corresponding to y , we define a random tilt of ν_N , which we denote by ν_N^Γ , as the measure

$$\nu_N^\Gamma = T(\sigma) d\nu_N \quad (3.3)$$

where T is given by

$$T(\sigma) = \exp \left(\log(\cosh(y(\sigma))) - \log \left(\int_{\Sigma_N} \cosh(y(\sigma)) d\nu_N \right) \right). \quad (3.4)$$

Observe that since $\cosh(x) \geq 1$, these measures are mutually absolutely continuous.

Assume that for H' , the limiting overlap distribution satisfies $\zeta(q_*) > 0$. As ξ is generic, there is a collection of sets, $\{X_{\alpha, N}\} \subset \Sigma_N$, that satisfies items 1.-5. of Theorem A.1, with respect to the

measure ν_N . We drop the N dependence in the notation of $X_{\alpha,N}$ and write X_α . For each $\alpha \in \mathbb{N}$, we define the measure

$$\nu_{\alpha,N} = \nu_N(\cdot | X_\alpha),$$

when X_α is non-null, and on the event that it is null, let this be $\delta_{(1,\dots,1)}$. Finally we let $\nu_{\alpha,N}^\top$ be the measure on $\{-1, 1\} \times \mathbb{R} \times \Sigma_N$ such that for ϕ continuous and bounded,

$$\int \phi(s, y, \sigma) \nu_{\alpha,N}^\top = \frac{\int_{X_\alpha} \int_{\Sigma_1} \phi(s, y, \sigma) e^{sy(\sigma)} ds d\nu_N(\sigma)}{\int_{X_\alpha} 2 \cosh(y) d\nu_N}. \quad (3.5)$$

For the purposes of this section, let $\langle \cdot \rangle_{\alpha,N}$ denote integration with respect to $\nu_{\alpha,N}$, and $\langle \cdot \rangle_{\alpha,N}^\top$ to denote integration with respect to $\nu_{\alpha,N}^\top$.

Let $((s_{\alpha,N}^i, y_{\alpha,N}^i))_{i \geq 1}$ then be i.i.d. draws from $\nu_{\alpha,N}^\top$. Here, we marginalize over the last coordinate σ^i . The goal of this section is to study the convergence of the joint law this pair. In particular, let $\sigma \in \mathcal{H}$ have $\|\sigma\| = \sqrt{q_*}$ and let μ_σ be the state measure at σ with boundary data $f = Y_{q_*}^\sigma$, as in Section 1.2. Finally, let $((s^i, y^i))_i$ be i.i.d. draws from μ_σ . The main theorem of this section is the following.

Theorem 3.1. *Assume that for H' , the limiting overlap distribution satisfies $\zeta(q_*) > 0$. For each $\alpha \in \mathbb{N}$,*

$$((s_{\alpha,N}^i, y_{\alpha,N}^i))_i \rightarrow ((s^i, y^i))_i$$

in distribution.

Recall now that (s^i, y^i) satisfies the conclusion of Theorem 1.9. (As noted in Section 1.2, the equation in Theorem 1.9 is in fact independent of the choice of boundary data.) As a consequence, we have the following corollary.

Corollary 3.2. *In the setting of Theorem 3.1, we have that*

$$\left(\langle s \rangle_{\alpha,N}^\top - \tanh \left(\langle y \rangle_{\alpha,N}^\top - (\xi'(1) - \xi'(q_*)) \langle s \rangle_{\alpha,N}^\top \right) \right)_{\alpha \in \mathbb{N}} \rightarrow 0$$

in distribution.

The goal of this section is to prove these two results. We begin by proving that the overlap distribution for $\nu_{\alpha,N}$ has a simple limit. We then prove Portmanteau type theorems for $(s_{\alpha,N}, y_{\alpha,N})$. These results allow us to conclude that statistics of $(s_{\alpha,N}, y_{\alpha,N})$ are a continuous functionals of the overlap distribution of $\nu_{\alpha,N}$ (not $\nu_{\alpha,N}^\top$). Since the latter converges, we then conclude Theorem 3.1. The proof of Corollary 3.2 is then an application of this theorem combined with Theorem 1.9.

3.1 Convergence of Overlaps between clusters.

We now prove that the $\nu_{\alpha,N}$ are replica symmetric. Fix $\alpha \in \mathbb{N}$. Let $(\sigma_i^\alpha)_{i=1}^\infty$ be drawn from $\nu_{\alpha,N}^{\otimes \infty}$ and consider R_N to be the doubly infinite overlap array defined by

$$R_N = (R(\sigma_i^\alpha, \sigma_j^\alpha)).$$

Finally, let Q be the deterministic matrix which is doubly infinite, all 1 on the diagonal and q_* on the off-diagonal. We then have the following theorem.

Theorem 3.3. *We have that*

$$R_N \xrightarrow{(d)} Q.$$

Proof. By standard properties of product spaces, it suffices to show that for any k ,

$$\mathbb{E} \int_{X_\alpha^k} F(R_N^k) d\nu_{\alpha,N} \rightarrow F(Q^k). \quad (3.6)$$

Here F is some smooth function on $[-1, 1]^{k^2}$ and by R_N^k and Q^k are the overlap matrix for k i.i.d. draws from $\nu_{\alpha,N}$ and the first k -by- k entries of Q respectively. It suffices to work on the event that X_α is non-empty. Since F is smooth, observe that it suffices to show that

$$\mathbb{E} \int_{X_\alpha^k} \|R_N^k - Q^k\|_1 d\nu_{\alpha,N}^{\otimes k} = o_N(1).$$

To this end, observe that

$$\int_{X_\alpha^k} \|R_N^k - Q^k\|_1 d\nu_{\alpha,N}^{\otimes k} = k \cdot (k-1) \frac{\int_{X_\alpha^2} |R_{12} - q_*| d\nu_{\alpha,N}^{\otimes 2}}{\nu(X_\alpha)^2},$$

where R_{12} is the overlap of two replica from $\nu_{\alpha,N}$ and the diagonal terms cancelled. This goes to zero in probability by Theorem A.1 items 4 and 5. \square

3.2 Continuity and Portmanteau-type results

We now collect some continuity and Portmanteau type theorems which will be useful in the following.

Lemma 3.4. *For each α , the convergence*

$$((s_{\alpha,N}^i, y_{\alpha,N}^i)) \xrightarrow{(d)} ((s^i, y^i))$$

holds if and only if for every k , $d : [k] \rightarrow \{0, 1\}$, and family of continuous bounded functions $\{\phi_i\}$,

$$\mathbb{E} \prod_{i \in [k]} (s_{\alpha,N}^i)^{d(i)} \phi_i(y_{\alpha,N}^i) \rightarrow \mathbb{E} \prod_{i \in [k]} (s^i)^{d(i)} \phi_i(y^i). \quad (3.7)$$

Furthermore, it is necessary and sufficient to take ϕ of polynomial growth.

This result is a standard consequence of the fact that $s_{\alpha,N}$ are $\{\pm 1\}$ valued and $\{y_{\alpha,N}\}$ have uniformly bounded sub-Gaussian tails (see Lemma A.5), so we omit its proof.

Finally we note the following continuity result which is a consequence of Theorem 3.3.

Lemma 3.5. *For any $k, \ell \geq 1$ and any family of continuous bounded functions $\{\phi_i\}_{i \in [\ell]}$, we have that*

$$\begin{aligned} & \mathbb{E} \int_{X_\alpha^{k+\ell}} \prod_{i \in [\ell]} \phi_i(y(\sigma_i^\alpha)) \prod_{j=\ell+1}^{\ell+k} \cosh(y(\sigma_j^\alpha)) d\nu_{\alpha,N}^{\otimes \ell+k} \\ & \rightarrow \mathbb{E} \left[\prod_{i \in [\ell]} \mathbb{E}(\phi_i(Y_1) | Y_{q_*} = Y_{q_*}(\sigma)) \cdot \mathbb{E}(\cosh(Y_1) | Y_{q_*} = Y_{q_*}(\sigma))^k \right] \end{aligned} \quad (3.8)$$

where Y_t is the driving process and σ is any $\sigma \in \mathcal{H}$ such that $\|\sigma\|^2 = q_$.*

Proof. Recall from (3.2), that y is independent of H' by construction. Thus it is independent of $\nu_{\alpha,N}$ and X_α . We may then integrate the lefthand side of (3.8) first in y , to obtain

$$\mathbb{E} \int_{X_\alpha^{k+\ell}} F(\xi'(R) + o_N(1)) d\nu_{\alpha,N}^{\otimes k+\ell}$$

for some continuous bounded function F , where by $\xi'(R) + o_N(1)$, we mean that we apply a function of this type to R entry wise. (The fact that such an F exists is a standard argument, see for example [18].) By a mollification argument, it suffices to study the convergence of

$$\mathbb{E} \int_{X_\alpha^{k+\ell}} F(\xi'(R)) d\nu_{\alpha,N}^{\otimes k+\ell}$$

where this is the same function F as above. By Theorem 3.3, this converges to $F(\xi'(Q))$. It remains to understand $F(\xi'(Q))$. By definition

$$F(\xi'(Q)) = \mathbb{E} \left[\left(\prod_{i \in [\ell]} \mathbb{E}(\phi_i(Y_1) | Y_{q_*} = Y_{q_*}(\sigma)) \right) \mathbb{E}(\cosh(Y_1) | Y_{q_*} = Y_{q_*}(\sigma))^k \right],$$

as desired. \square

3.3 Proofs of main theorems

We can now turn to the proofs of the main results.

Proof of Theorem 3.1. Fix α . It suffices to work on the event that X_α is non-empty. By Lemma 3.4, it suffices to prove (3.7) for each n , $d : [n] \rightarrow \{0, 1\}$ and family of continuous bounded $\{\phi_i\}$. Furthermore, we claim that it suffices to prove

$$\mathbb{E} \prod_{i \in [n]} \langle \phi_i(y) \rangle_{\alpha,N}^\top \rightarrow \mathbb{E} \prod_{i \in [n]} \phi_i(y_\alpha^i). \quad (3.9)$$

To see this, simply note that

$$\begin{aligned} \mathbb{E} \prod_i (s_{\alpha,N}^i)^{d(i)} \phi_i(y_{\alpha,N}) &= \mathbb{E} \prod_i \left\langle s^{d(i)} \phi_i(y) \right\rangle_{\alpha,N}^\top \\ &= \mathbb{E} \prod_i \frac{\int_{X_\alpha} \int_{\Sigma_1} \phi_i(y(\sigma)) s^{d(i)} e^{sy(\sigma)} ds d\nu_N}{\int_{X_\alpha} 2 \cosh(y) d\nu_N} \\ &= \mathbb{E} \prod_i \langle f_{d(i)}(y) \phi_i(y) \rangle_{\alpha,N}^\top, \end{aligned}$$

where $f_d(x) = \tanh(x)$ if $d = 1$ and 1 if $d = 0$.

With this claim in hand, we now prove (3.9). To this end, fix ϕ_i as above. By (3.5),

$$\mathbb{E} \prod_i \phi_i(y_{\alpha,N}^i) = \mathbb{E} \prod_i \frac{\int_{X_\alpha} \phi_i(y(\sigma)) \cosh(y(\sigma)) d\nu_N}{\int_{X_\alpha} \cosh(y(\sigma)) d\nu_N}.$$

Observe that $Z_\alpha = \int_{X_\alpha} \cosh(y) d\nu_N$ satisfies $Z_\alpha \geq 1$. By Lemma A.5,

$$P(Z_\alpha \geq L) \lesssim_\xi 1/L$$

uniformly in N . Thus by a standard approximation argument, we can approximate $1/Z_\alpha^n$ by polynomials in Z_α in the above expectations. In particular, it suffices to study limits of integrals of the form

$$\mathbb{E} \prod_i \int_{X_\alpha} \phi_i(y(\sigma)) \cosh(y(\sigma)) d\nu_N \cdot \left(\int_{X_\alpha} \cosh(y(\sigma)) d\nu_N \right)^l.$$

This is exactly of the form (3.8) with $k = l$, $\ell = 1$ and the family $\{\phi_i(y) \cdot \cosh(y)\}_{i \in [n]}$ by Fubini's theorem. Thus by Lemma 3.5,

$$\mathbb{E} \phi(y_{\alpha, N}) \rightarrow \mathbb{E} \prod_i \frac{\mathbb{E}(\phi_i(Y_1) \cosh(Y_1) | Y_{q_*} = Y_{q_*}(\sigma))}{\mathbb{E}(\cosh(Y_1) | Y_{q_*} = Y_{q_*}(\sigma))},$$

where σ is any $\sigma \in \mathcal{H}$ such that $\|\sigma\|^2 = q_*$.

It remains to recognize the righthand side of the above display as a state average. By Girsanov's theorem and properties of the Parisi IVP (1.13) (see for instance [11, Lemma 8.3.1]), we have that

$$\begin{aligned} \frac{\mathbb{E}(\phi(Y_1) \cosh(Y_1) | Y_{q_*} = Y_{q_*}(\sigma))}{\mathbb{E}(\cosh(Y_1) | Y_{q_*} = Y_{q_*}(\sigma))} &= \mathbb{E} \left(\phi(Y_1) e^{u(1, Y_1) - u(q_*, Y_{q_*}) - \frac{1}{2}(\xi'(1) - \xi'(q_*))} | Y_{q_*} = Y_{q_*}(\sigma) \right) \\ &= \mathbb{E} \langle \phi(y) \rangle_\alpha^Y, \end{aligned}$$

where u is the solution of (1.13) corresponding to ζ and the last equality is by definition. Thus

$$\mathbb{E} \prod \phi_i(y_{\alpha, N}) \rightarrow \mathbb{E} \prod \langle \phi_i(y) \rangle_\alpha^Y = \mathbb{E} \prod \phi_i(y_\alpha),$$

as desired. \square

Proof of Corollary 3.2. Let $m_{\alpha, N} = \langle s \rangle_{\alpha, N}$ and $h_{\alpha, N} = \langle y \rangle_{\alpha, N}$. We claim that for each $\alpha \in \mathbb{N}$,

$$(m_{\alpha, N}, h_{\alpha, N}) \xrightarrow{(d)} (m_\alpha, h_\alpha).$$

Suppose first that this claim is true. Then the result immediately follows from Theorem 1.9.

We now turn to the claim. Observe that by Lemma A.5, these random variables have sub-Gaussian tails. Thus it suffices to prove convergence of the moments

$$\mathbb{E} m_{\alpha, N}^{k_1} h_{\alpha, N}^{k_2}.$$

To this end, let $k = k_1 + k_2$ and let $\{\psi_j\}_{j \in [k]}$ satisfy $\psi_j = 1$ if $j \leq k_1$ and $\psi_j(x) = x$ if $j > k_1$. Finally let $d : [k] \rightarrow \{0, 1\}$ be such that $d(i) = 1$ if $i \leq k_1$ and $d(i) = 0$ otherwise. Then, by Lemma 3.4 and Theorem 3.1, we have that

$$\mathbb{E} m_{\alpha, N}^{k_1} h_{\alpha, N}^{k_2} = \mathbb{E} \prod_j (s_{\alpha, N}^j)^{d(j)} \psi_j(y_{\alpha, N}^j) \rightarrow \mathbb{E} \prod_j (s_\alpha^j)^{d(j)} \psi_j(y_\alpha^j) = \mathbb{E} m_\alpha^{k_1} h_\alpha^{k_2}$$

as desired. \square

4 Stability of clusters under lifts

In this section, we show that important properties of the pure states are carried over after lifting in one coordinate. We start with the following construction. For $\sigma = (\sigma_1, \dots, \sigma_N) \in \Sigma_N$, let $\rho(\sigma) = (\sigma_2, \dots, \sigma_N) \in \Sigma_{N-1}$. For any mixed p -spin glass model, the Hamiltonian, H_N , decomposes into a sum of three Gaussian processes:

$$H_N(\sigma) = \tilde{H}_N(\rho(\sigma)) + \sigma_1 y_N(\rho(\sigma)) + r_N(\sigma_1, \rho(\sigma)). \quad (4.1)$$

Properties of these Gaussian processes are described in Lemma A.6. For $\sigma \in \Sigma_{N-1}$ set

$$H'_N(\sigma) := \tilde{H}_N(\sigma) + r_N(1, \sigma)$$

and let μ'_N be the Gibbs measure corresponding to the Hamiltonian H'_N . This Hamiltonian, and thus μ'_N , is independent of y_N . We are thus in the setting of Section 3 where H'_N satisfies (3.1) and y_N satisfies (3.2).

Let $\tilde{W}_{\alpha, N-1}$, $\alpha \in \mathbb{N}$ be the subsets of Σ_{N-1} constructed via Theorem A.1 relative to the measure μ'_N . Set

$$W_{\alpha, N}^\dagger = \Sigma_1 \times \tilde{W}_{\alpha, N-1} \subset \Sigma_N. \quad (4.2)$$

Order the sets $W_{\alpha, N}^\dagger$ with respect to their μ_N masses. That is, define subsets $W_{\alpha, N} \subset \Sigma_N$, for $\alpha \in \mathbb{N}$, such that

$$\mu_N(W_{1, N}) \geq \mu_N(W_{2, N}) \geq \dots \quad (4.3)$$

and so that

$$W_{\alpha, N} = W_{\pi_N(\beta), N}^\dagger,$$

for some (random) automorphism $\pi_N : \mathbb{N} \rightarrow \mathbb{N}$.

Remark 4.1. Note that there is not a unique way to define the projection π_N since, there are possibly ties $W_\alpha = W_\beta$. Note, however, this only introduces a finite indeterminacy as there are only finitely many such sets that are non-empty by construction. The reader can take any tie breaking rule.

The goal of this section is to show that the collection $(W_{\alpha, N})_{\alpha \in \mathbb{N}}$ also satisfies items 1.-5. from Theorem A.1. (For the rest of section, we drop the subscript N of our notation.) The main idea is that at the level of overlaps, the measure μ on the sets W_α will essentially be the same as the measure $(\mu')^\top$ on the sets $\tilde{W}_{\pi(\alpha)}$. Since on Σ_{N-1} , $(\mu')^\top \ll \mu'$, overlap events that are rare for μ' will still be rare for $(\mu')^\top$. We begin by recording the following lemma which is a quantification of this observation.

Recall the local field $y = y_N$ from (1.5) and the function T from (3.4). Let

$$\tilde{K}(\mu') = \left(\int \cosh(2y) d\mu' \right)^{1/2}. \quad (4.4)$$

Lemma 4.2 (Tilting Lemma). *There are constants $C, c > 0$ such that with probability at least $1 - \frac{1}{c}e^{-cN}$,*

$$\left(1 - \frac{C}{\sqrt{N}}\right) \int_A T d\mu' \leq \mu(\Sigma_1 \times A) \leq \left(1 + \frac{C}{\sqrt{N}}\right) \int_A T d\mu', \quad \forall A \subset \Sigma_{N-1}.$$

In particular,

$$\mu(\Sigma_1 \times A) \leq \tilde{K}(\mu') \left(1 + \frac{C}{\sqrt{N}}\right) \sqrt{\mu'(A)}.$$

Proof. This result immediately follows from Lemma A.6. Observe that if we let

$$\Delta = 2 \max_{\sigma \in \Sigma_{N-1}} |r(1, \sigma) - r(-1, \sigma)|,$$

then

$$\mu(\Sigma_1 \times A) = \frac{\int_A \int_{\Sigma_1} e^{\tilde{H}(\sigma) + \epsilon y(\sigma) + r(\epsilon, \sigma)} d\epsilon d\sigma}{\int_{\Sigma_{N-1}} \int_{\Sigma_1} e^{\tilde{H}(\sigma) + \epsilon y(\sigma) + r(\epsilon, \sigma)} d\epsilon d\sigma} \leq \int_A T(\sigma) d\mu' e^\Delta.$$

Similarly

$$\mu(\Sigma_1 \times A) \geq \int_A T(\sigma) d\mu' e^{-\Delta}.$$

The first result then follows by Lemma A.6, and the second result follows from the first and the Cauchy-Schwarz inequality. \square

We now start by proving the properties mentioned above.

Lemma 4.3. *Let $q'_N = q_{N-1}$, $a'_N = a_{N-1}$, $b'_N = b_N^{1/4}$, and $\epsilon'_N = \epsilon_N^{1/4}$. Then the sets $\{W_\alpha\}_{\alpha \in [m_N]}$ satisfy items 1 – 4 Theorem A.1 with probability $1 - o_N(1)$, where the sequences $q'_N, a'_N, b'_N, \epsilon'_N$ and m_N satisfy those conditions, except with $m_N \cdot b_N^{1/2} \rightarrow 0$.*

Proof. Since the sets \tilde{W}_α are disjoint, W_α^\dagger and W_α are as well and satisfy

$$(\cup_\alpha W_\alpha)^c = (\cup_\alpha W_\alpha^\dagger)^c = \left(\cup_\alpha \Sigma_1 \times \tilde{W}_\alpha \right)^c = \Sigma_1 \times \left(\cup_\alpha \tilde{W}_\alpha \right)^c.$$

Thus by the Tilting Lemma (Lemma 4.2) and item 1 of Theorem A.1, we have that with high probability,

$$\mu((\cup_\alpha W_\alpha)^c) \leq \left(1 + \frac{C}{\sqrt{N}} \right) \tilde{K}(\mu') \cdot \sqrt{\epsilon_N}. \quad (4.5)$$

Furthermore, by the Tilting Lemma and item 2 of Theorem A.1, we obtain for $\beta = \pi^{-1}(\alpha)$

$$\begin{aligned} \mu^{\otimes 2}(\sigma^1, \sigma^2 \in W_\alpha : R_{12} \leq q_{N-1} - 2a_{N-1}) \\ \leq \tilde{K}(\mu')^2 \left(1 + \frac{C}{\sqrt{N}} \right) \sqrt{(\mu')^{\otimes 2} \left(\sigma^1, \sigma^2 \in \tilde{W}_\beta : R_{12} \leq q_{N-1} - 2a_{N-1} + \frac{1}{N} \right)} \\ \leq \tilde{K}(\mu')^2 \left(1 + \frac{C}{\sqrt{N}} \right) \sqrt{b_N}, \end{aligned}$$

where we used the fact that we may take $a_{N-1} \geq \frac{1}{N}$. Argue similarly to get that for $\alpha_1 \neq \alpha_2$,

$$\mu^{\otimes 2}(\sigma^1 \in W_{\alpha_1}, \sigma^2 \in W_{\alpha_2} : R_{12} \geq q_{N-1} + 2a_{N-1}) \leq \tilde{K}(\mu')^2 \left(1 + \frac{C}{\sqrt{N}} \right) \sqrt{b_N}.$$

Observe that by Lemma A.5, with probability tending to 1, $\tilde{K}(\mu') \leq b_N^{-1/8} \vee \epsilon_N^{-1/4}$. This yields the desired result after observing that since $\zeta_N[q_N + a_N, 1] \geq \zeta\{q_*\} - b_N$, for N sufficiently large, the same is true for q'_N, a'_N and b'_N , and that item 4 in Theorem A.1 is implied by this fact regarding ζ_N and items 2 and 3. \square

It remains to show that the weights $\mu(W_\alpha)$ converge to a Poisson-Dirichlet process.

Lemma 4.4. *We have that*

$$(\mu(W_\alpha))_{\alpha \in \mathbb{N}} \rightarrow (v_\alpha)_{\alpha \in \mathbb{N}}$$

in distribution on the space of mass partitions \mathcal{P}_m .

Proof. Recall that $\{\mu_N\}$ satisfy the approximate Ghirlanda-Guerra identities since H_N is a generic model. Let $U_{12} = U(\sigma^1, \sigma^2)$ be

$$U_{12} = \mathbb{1} \{ \exists \alpha \in \mathbb{N} : \sigma^1, \sigma^2 \in W_\alpha \}$$

and let $L_N = \{\sigma^1, \sigma^2 \in \cup_\alpha W_\alpha\}$. Then by the arguments of [10, Section 6], in order to prove that this sequence converges, it suffices to prove that for some $\phi_{\kappa, \lambda}$ which satisfies

$$\phi_{\kappa, \lambda}(x) = \begin{cases} 0 & x \leq q_* - \kappa \\ 1 & x \geq q_* - \lambda \end{cases}$$

and interpolates between the two values for $x \in [q_* - \kappa, q_* - \lambda]$, we have

$$\lim_{\kappa, \lambda \rightarrow 0} \overline{\lim}_{N \rightarrow \infty} \mathbb{E} \langle |U_{12} - \phi_{\kappa, \lambda}| \rangle = 0.$$

To see this, if we denote $|U_{12} - \phi_{\kappa, \lambda}| = \clubsuit$, then

$$\mathbb{E} \langle \clubsuit \rangle_\mu \leq \mathbb{E} \langle \clubsuit \mathbb{1} \{L_N\} \rangle + o_N(1)$$

where the fact that the second term is $o_N(1)$ follows from (4.5). Now

$$\begin{aligned} \mathbb{E} \langle \clubsuit L_N \rangle &= \mathbb{E} \langle \clubsuit \mathbb{1} \{L_N, R_{12} \geq q_* - \lambda\} U_{12} \rangle + \mathbb{E} \langle \clubsuit \mathbb{1} \{L_N, R_{12} \leq q_* - \lambda\} U_{12} \rangle \\ &\quad + \mathbb{E} \langle \clubsuit \mathbb{1} \{L_N, R_{12} \geq q_* - \kappa\} (1 - U_{12}) \rangle + \mathbb{E} \langle \clubsuit \mathbb{1} \{L_N, R_{12} \leq q_* - \kappa\} (1 - U_{12}) \rangle \\ &= I + II + III + IV. \end{aligned}$$

Note that $I = IV = 0$ identically. It remains to estimate II and III .

We start with II . Observe that

$$II \leq 2\mathbb{E} \langle U_{12} (\mathbb{1} \{R_{12} \leq q_* - 2a_{N-1}\}) \rangle$$

for N large enough, which is bounded by b'_N by Lemma 4.3.

Now to estimate III . Note that for N sufficiently large,

$$\mathbb{E} \langle \clubsuit (1 - U_{12}) (\mathbb{1} \{R_{12} \geq q_{N-1} + 2a_{N-1}\} + \mathbb{1} \{R_{12} \in [q_* - \kappa, q_N + 2a_{N-1}]\}) \rangle_\mu \leq b'_N + (a).$$

By the tilting lemma,

$$(a) \leq \|\tilde{K}\|_4^2 \cdot (\mathbb{E} \mu^{\otimes 2}[q_* - 2\kappa, q_{N-1} + a_{N-1}])^{1/2}.$$

By the choice of q_N and a_N (see the first display in Theorem A.1), we have that

$$\overline{\lim} \zeta_N[q_* - 2\kappa, q_{N-1} + a_{N-1}] = \overline{\lim} (\zeta_N[q_* - 2\kappa, 1] - \zeta_N[q_{N-1} + a_{N-1}, 1]) = 0.$$

Thus combining these estimates and Lemma A.5 we see that sending $N \rightarrow \infty$, $\lambda \rightarrow 0$ and then $\kappa \rightarrow 0$ yields the result. \square

5 Essential uniqueness of clusters

In this section, we show that sets that satisfy the properties from Theorem A.1 with respect to μ are asymptotically unique.

Let $\{C_\alpha\}$ be constructed as in Theorem A.1 for the measure μ_N . Recall that they are labelled in decreasing order, i.e.,

$$\mu_N(C_\alpha) \geq \mu_N(C_{\alpha+1}).$$

Let $a_N, b_N, m_N, q_N \rightarrow q_*$, and ϵ_N be as in that theorem. Let $\{X_\alpha\}_{\alpha \in [m_N]}$ be another collection of sets that satisfies items 1-5 of Theorem A.1, with constants q'_N, a'_N, b'_N and ϵ'_N as in that theorem except with $m_N \cdot b'_N \rightarrow 0$.

The main goal of this section is to prove that, the pure states C_α and the sets X_α are effectively the same, as far as μ is concerned.

Theorem 5.1 (Essential uniqueness). *Suppose that we have*

$$\zeta_N [(q'_N - a'_N), (q'_N + a'_N)] + \zeta_N [(q_N - a_N), (q_N + a_N)] \rightarrow 0. \quad (5.1)$$

Then, for each $\alpha \in \mathbb{N}$, we have that

$$\mu_N(C_\alpha \Delta X_\alpha) \rightarrow 0 \quad (5.2)$$

in probability, where Δ denotes the symmetric difference.

As a corollary of this we get the following.

Corollary 5.2. *Let W_α be as in Lemma 4.3. Then (5.1) holds. In particular,*

$$\mu_N(C_\alpha \Delta W_\alpha) \rightarrow 0 \quad (5.3)$$

in probability.

Proof. This follows by Lemma 4.3 and Theorem 5.1 after recalling that

$$\begin{aligned} \zeta_N[q'_N + a'_N, 1] &\geq \zeta[q_*] - o_N(1) \\ \zeta_N[q_N + a_N, 1] &\geq \zeta[q_*] - o_N(1). \end{aligned}$$

Indeed, this implies that

$$\zeta_N[q'_N - a'_N, q_N + a_N] = \zeta_N[q_{N-1} - a_{N-1}, q_N + a_N] \rightarrow 0.$$

The same argument holds for the second limit. □

The idea of the proof Theorem 5.1 is that the overlap properties of the sets (X_α) and (C_α) will imply that each of the first n (X_α) 's will be supported by one the first M (C_α) 's for some M large but fixed, and *vice versa*. The ranking of the states and basic properties of the Poisson-Dirichlet process will then imply that, in fact, for each α , the sets X_α and C_α are actually supported by each other.

For this we will need the following three lemmas. Their proofs are deferred to the end of this section and follow from properties of the Poisson-Dirichlet process. The first lemma says that there is not much mass in the the tail of the collections X_α and C_α .

Lemma 5.3. *For every $\epsilon > 0$, there is an $N_0(\epsilon)$ and $M(\epsilon)$ such that if*

$$E_N(\epsilon) = \left\{ \mu_N \left(\cup_{\alpha \geq M(\epsilon)} X_\alpha \right) > \frac{\epsilon}{2} \right\} \cup \left\{ \mu_N \left(\cup_{\alpha \geq M(\epsilon)} C_\alpha \right) > \frac{\epsilon}{2} \right\}$$

then for $N \geq N_0(\epsilon)$,

$$\mathbb{P}[E_N(\epsilon)] \leq \epsilon.$$

The second lemma says that, for any fixed n , the first n states (C_k) and (X_k) must have non-negligible μ_N mass as N goes to infinity.

Lemma 5.4. *Fix $n \geq 1$ and $\delta > 0$. Let $F_N(n, \delta)$ be the event that*

$$\begin{aligned} \mu_N(X_1) &> \dots > \mu_N(X_n) > \delta \\ \mu_N(C_1) &> \dots > \mu_N(C_n) > \delta, \end{aligned}$$

then there is a function $f_{1,n}$ satisfying $\lim_{\delta \rightarrow 0} f_{1,n}(\delta) = 0$ and an $N_1(n, \delta)$ such that for $N \geq N_1(n, \delta)$,

$$\mathbb{P}[F_N(n, \delta)] \geq 1 - f_1(\delta).$$

The last lemma concerns the gap between the masses of states.

Lemma 5.5. *Fix $\eta > 0$ and $n \geq 1$. Let*

$$\begin{aligned} I_N(\eta, n) &= \{ \mu_N(C_i) - \mu_N(C_{i+1}) > \eta \ \forall i \in [n-1] \} \\ &\cap \{ \mu_N(X_i) - \mu_N(X_{i+1}) > \eta \ \forall i \in [n-1] \}. \end{aligned}$$

Then there is a function $f_2(\eta, n)$ and an $N_2(\eta, n)$, such that for $N \geq N_2(\eta, n)$,

$$\mathbb{P}(I_N(\eta, n)) \geq 1 - f_2(\eta, n),$$

where for each n , $f_2(\eta, n) \rightarrow 0$ as $\eta \rightarrow 0$.

Given $\varepsilon > 0$, choose δ , ϵ , and η by combining Lemma 5.3-5.5, such that if

$$\mathcal{E}_N(\epsilon, \delta, n, \eta) := E_N^c(\epsilon) \cap F_N(n, \delta) \cap I_N(\eta, n) \cap J_N,$$

where J_N is the event that the conclusions of Theorem A.1 hold then

$$\mathbb{P}[\mathcal{E}_N] > 1 - \varepsilon, \tag{5.4}$$

for all $N \geq N_0(\varepsilon)$.

Proof of Theorem 5.1. We want to show that for each $\rho > 0, \varepsilon > 0$ and α ,

$$\mathbb{P}(\mu_N(C_\alpha \Delta X_\alpha) > \rho) \leq \varepsilon. \tag{5.5}$$

Fix ρ, ε , and α . Let $n > \alpha$. Let $N \geq N_0(\varepsilon/2)$ where N_0 is defined as in (5.4). By (5.1) and Markov's inequality, there is a $c_N \rightarrow 0$ such that with probability $1 - o_N(1)$,

$$\mu_N^{\otimes 2}(R_{12} \in [q'_N - a'_N, q_N + a_N]) \leq c_N. \tag{5.6}$$

Choose N sufficiently large that

$$\frac{2M(\epsilon)}{\delta}(m_N(b_N + b'_N) + c_N) + \epsilon_N < \frac{\rho \wedge \eta \wedge \varepsilon}{2}$$

where δ, ϵ, η are defined as above with $\delta = \epsilon$ and depend only on ε . We can do this since by assumption,

$$(b'_N + b_N) \cdot m_N = o_N(1).$$

We will prove shortly that on \mathcal{E}_N , for

$$\iota_N = \frac{2M(\epsilon)}{\epsilon}(m_N(b_N + b'_N) + c_N) + \epsilon_N,$$

we have that

$$\begin{aligned} \mu_N(C_\alpha \setminus X_\alpha) &\leq \iota_N, \\ \mu_N(X_\alpha \setminus C_\alpha) &\leq \iota_N. \end{aligned} \tag{5.7}$$

Note that (5.7) immediately implies (5.5) as desired. \square

Proof of (5.7). We begin by defining two maps $\pi_1, \pi_2 : [n] \rightarrow [M(\epsilon)]$. On the event \mathcal{E}_N , for each i , we let $\pi_1(i)$ be the first $j \in [M(\epsilon)]$ such that

$$\mu_N(X_i \cap C_{\pi_1(i)}) \geq \frac{\epsilon}{2 \cdot M(\epsilon)}$$

holds and let $\pi_2(i)$ be the first $j \in [M(\epsilon)]$ such that

$$\mu_N(X_{\pi_2(i)} \cap C_i) \geq \frac{\epsilon}{2 \cdot M(\epsilon)}$$

holds. That such j exist follows by definition of \mathcal{E}_N . On \mathcal{E}_N^c , let $\pi_1 = \pi_2 = Id$. This provides two random maps $\pi_i : [n] \rightarrow [M(\epsilon)]$, $i = 1, 2$.

Suppose for the moment that on \mathcal{E}_N ,

$$\begin{aligned} \mu_N(X_i \cap C_{\pi_1(i)}) &\geq \mu_N(X_i) - \iota_N \\ \mu_N(C_i \cap X_{\pi_2(i)}) &\geq \mu_N(C_i) - \iota_N. \end{aligned} \tag{5.8}$$

The inequality, (5.7), will then follow provided we can show that $\pi_1 = \pi_2 = Id$ on \mathcal{E}_N . Let us first show that these maps are the identity given (5.8). We then prove (5.8).

The proof that these maps are the identity map is by induction. Suppose first that $\pi_2(1) = 1$. If $\pi_1(1) > 1$, then by (5.8),

$$\begin{aligned} \mu_N(C_1) &\leq \mu_N(X_1) + \iota_N \\ &\leq \mu_N(C_{\pi_1(1)}) + 2\iota_N \leq \mu_N(C_2) + 2\iota_N. \end{aligned}$$

This implies that

$$\mu_N(C_1) - \mu_N(C_2) \leq 2\iota_N.$$

Since $\iota_N \rightarrow 0$, this contradicts the definition of \mathcal{E}_N . By symmetry, the same argument works if $\pi_1(1) = 1$ and $\pi_2(1) > 1$.

Now assume that $\pi_2(1) > 1$ and $\pi_1(1) > 1$. By the ordering of these sets,

$$\begin{aligned}\mu_N(C_1) &\leq \mu_N(X_{\pi_2(1)}) + \iota_N \leq \mu_N(X_1) + \iota_N \\ &\leq \mu_N(C_{\pi_1(1)}) + 2\iota_N.\end{aligned}$$

This is, again, a contradiction. Thus $\pi_1(1) = 1 = \pi_2(1)$.

Assume now that $\pi_1(i) = \pi_2(i) = i$ for all $i \in [k-1]$. By the same reasoning as in the base case, if $\pi_2(k) \neq k$, then it must be that $\pi_2(k) < k$. This, however, implies that

$$\mu_N(C_k) \leq \mu_N(X_{\pi_2(k)} \setminus C_{\pi_2(k)}) + \iota_N.$$

But

$$\begin{aligned}\mu_N(X_{\pi_2(k)} \setminus C_{\pi_2(k)}) &\leq \mu_N(X_{\pi_2(k)}) - \mu_N(X_{\pi_2(k)} \cap C_{\pi_2(k)}) \\ &\leq \iota_N,\end{aligned}$$

where we used the induction hypothesis in the last inequality. This implies that eventually $\mu C_\alpha \leq 2\iota_N$. This is, again, a contradiction since on \mathcal{E}_N , $\mu C_\alpha > \epsilon$. Thus, assuming (5.8), we have that $\pi_1 = \pi_2 = Id$ by induction.

We now prove (5.8) on the event \mathcal{E}_N . Fix $\alpha \in [n]$. We know that on this event,

$$\mu_N(X_\alpha \cap C_{\pi_1(\alpha)}) \geq \frac{\epsilon}{2M(\epsilon)}.$$

Now let $\ell \neq \pi_1(\alpha)$. Write

$$\begin{aligned}\mu_N(X_\alpha \cap C_\ell) &= \frac{1}{\mu_N(C_{\pi_1(\alpha)} \cap X_\alpha)} \cdot \mu_N(C_{\pi_1(\alpha)} \cap X_\alpha) \mu_N(X_\alpha \cap C_\ell) \\ &\leq \frac{1}{\mu_N(C_{\pi_1(\alpha)} \cap X_\alpha)} \mu_N^{\otimes 2}(\sigma^1 \in C_{\pi_1(\alpha)} \cap X_\alpha, \sigma^2 \in C_\ell \cap X_\alpha).\end{aligned}$$

Write the event $\{R_{12} \in [-1, 1]\}$ as

$$\begin{aligned}\{R_{12} \leq q'_N - a'_N\} \cup \{R_{12} \geq q_N + a_N\} \cup \{q'_N - a'_N < R_{12} < q_N + a_N\} \\ = I \cup II \cup III.\end{aligned}$$

Note that since we are in the event J_N ,

$$\mu_N^{\otimes 2}(\sigma^1 \in C_{\pi_1(\alpha)} \cap X_\alpha, \sigma^2 \in C_\ell \cap X_\alpha, I) \leq \mu_N^{\otimes 2}(\sigma^1, \sigma^2 \in X_\alpha, I) \leq b'_N,$$

while

$$\mu_N^{\otimes 2}(\sigma^1 \in C_{\pi_1(\alpha)} \cap X_\alpha, \sigma^2 \in C_\ell \cap X_\alpha, II) \leq b_N.$$

Summing on ℓ and using (5.6), we see that

$$\sum_{\ell \neq k} \mu_N(X_\alpha \cap C_\ell) \leq \frac{1}{\mu(C_{\pi_1(\alpha)} \cap X_\alpha)} (m_N(b'_N + b_N) + c_N) \leq \frac{2M}{\epsilon} \cdot (m_N(b'_N + b_N) + c_N).$$

This implies the first inequality of (5.8) after recalling that $\{C_\ell\}$ (almost) partitions Σ_N and that

$$\mu_N(X_\alpha \cap (\cup_\alpha C_\alpha)^c) \leq \epsilon_N,$$

by assumption. By symmetry, the same argument shows the second inequality holds as well. \square

5.1 Propositions regarding Poisson-Dirichlet.

We now turn to proving the above lemmas.

Proof of Lemma 5.3. Fix $\epsilon > 0$. Let (v_n) be $PD(1 - \zeta(q_*))$. Let $M(\epsilon)$ be such that

$$\mathbb{P} \left(\sum_{\alpha \geq M(\epsilon)} v_\alpha \geq \frac{\epsilon}{2} \right) \leq \frac{\epsilon}{4}.$$

Recall that $(\mu_N(X_\alpha)) \rightarrow (v_n)$ by Lemma 4.4. For (v_α^N) , this event is contained in the closed event (in the topology of mass partitions)

$$\left\{ \sum_{\alpha \leq M} v_\alpha^N \leq 1 - \epsilon/2 \right\},$$

and for v_α these events are equal. Thus we have that for N sufficiently large

$$\mathbb{P} \left(\mu_N \left(\bigcup_{\alpha > M(\epsilon)} X_\alpha \right) \geq \frac{\epsilon}{2} \right) \leq \frac{\epsilon}{2},$$

by the Portmanteau theorem. The same argument applies to the C_α . Intersecting these events yields the result by the inclusion-exclusion principle. \square

Proof of Lemma 5.4. It suffices to show that for $n \geq 1$ fixed,

$$\lim_{\epsilon \rightarrow 0} \lim_{N \rightarrow \infty} \mathbb{P}(\mu_N(X_1) > \dots > \mu_N(X_n) > \epsilon, \mu_N(C_1) > \dots > \mu_N(C_n) \geq \epsilon) = 1.$$

To see this, observe that by the inclusion-exclusion principle, and the Portmanteau lemma, we have that

$$\lim_{N \rightarrow \infty} \mathbb{P}(\mu_N(X_1) > \dots > \mu_N(X_n) > \epsilon, \mu_N(C_1) > \dots > \mu_N(C_n) > \epsilon) \geq 2\mathbb{P}(v_1 > \dots > v_n > \epsilon) - 1.$$

This goes to 1 as $\epsilon \rightarrow 0$ since $v_1 > \dots > v_n > 0$ almost surely. \square

Proof of Lemma 5.5. Again by the inclusion exclusion principle, it suffices to show that for every $i \in [n]$,

$$\lim_{\eta \rightarrow 0} \lim_{N \rightarrow \infty} \mathbb{P}(\mu_N(C_i) - \mu_N(C_{i+1}) > \eta) = 1.$$

Since $v_n - v_{n+1}$ is almost surely positive, this result is again immediate by the Portmanteau lemma. \square

6 TAP equation for a fixed coordinate

In this section we turn to the proof of Theorem 1.1. For the reader's convenience, let us briefly recap where we are and our plan of attack. Recall the construction of the states C_α from Theorem A.1 and the definition of $\langle \cdot \rangle_\alpha$. In Section 4 we constructed another collection of pure states $W_\alpha \subseteq \Sigma_N$ for the measure μ_N . As shown in Section 5, the sets C_α and W_α are essentially the same in each other.

The advantage of working with W_α lies in the fact that they are rearrangements of lifts of pure states of the measure μ'_{N-1} . This will allow us to avoid the first obstruction explained in Section 1.1: the measure μ'_{N-1} is now independent of the local field y_N . The rearrangement, however, is not independent of y_N . In particular the correlation between W_α and y is through the map π_N which takes W_α^\dagger to $W_{\pi(\alpha)}^\dagger = W_\alpha$.

To circumvent this obstruction we make the following observation. The measure μ conditioned on the set W_α^\dagger , is essentially the measure $(\mu')^\dagger$ conditioned on \tilde{W}_α . This will allow us to conclude that (1.7) holds by an application of Corollary 3.2, provided the rearrangement map π_N is not too wild. In particular, provided the map $\mu' \mapsto (\mu')^\dagger$ does not "charge the dust at infinity", the result will follow as a consequence of the following basic fact.

Lemma 6.1. *Let X^N be a sequence of $[-2, 2]^\mathbb{N}$ -valued random variables such that*

$$X^N \xrightarrow{(d)} 0.$$

Let π_N a sequence of S_∞ -valued random variables that satisfy the tightness criterion

$$\overline{\lim}_{M \rightarrow \infty} \overline{\lim}_{N \rightarrow \infty} P(\pi_N(n) \geq M) = 0 \quad \forall n. \quad (6.1)$$

Then if $Y^N = (X_{\pi(n)}^N)$, we have

$$Y^N \xrightarrow{(d)} 0.$$

Remark 6.2. Note that we made no assumptions on the joint law of π_N and X^N .

Proof. It suffices to prove this for each n individually. Observe that for each M ,

$$\mathbb{E} |Y_n^N| \leq \sum_{l=1}^M \mathbb{E} |X_l^N| + P(\pi_N(n) \geq M).$$

Sending $N \rightarrow \infty$ and then $M \rightarrow \infty$ then yields the result. \square

We begin this section by proving the tightness of the sequence π_N . The main result will then essentially be immediate, and is proved in the following subsection.

6.1 Tightness of the reshuffling

We begin this section by studying the random permutation $\pi : \mathbb{N} \rightarrow \mathbb{N}$ as defined in Section 4 by

$$W_{\alpha, N} = W_{\pi_N(\beta), N}^\dagger.$$

We recall its dependence on N by writing π_N instead of just π . We now show tightness for the sequence π_N .

Lemma 6.3 (Tightness). *We have that for each $n \in \mathbb{N}$,*

$$\overline{\lim}_{M \rightarrow \infty} \overline{\lim}_{N \rightarrow \infty} P(\pi_N^{-1}(n) \geq M) = 0.$$

Proof. Take N sufficiently large that $n \leq m_N$. Now observe that

$$\begin{aligned}
P(\pi_N^{-1}(n) \geq M) &= P(\exists k \geq M : \pi_N(k) = n) \\
&= P(\exists l \leq n, k \geq M : \mu(W_l^\dagger) \leq \mu(W_k^\dagger)) \\
&\leq \sum_{l=1}^n P(\exists k \geq M : \mu(W_l^\dagger) \leq \mu(W_k^\dagger)) \\
&\leq \sum_{l=1}^n P(\mu(W_l^\dagger) \leq \mu(\cup_{k \geq M} W_k^\dagger))
\end{aligned}$$

It thus suffices to prove this limit for each summand.

Now observe that for each $l \in [n]$ and each $\epsilon > 0$, the summand satisfies the inequality,

$$P(\mu(W_l^\dagger) \leq \mu(\cup_{k \geq M} W_k^\dagger)) \leq P(\mu(W_l^\dagger) \leq \epsilon) + P\left(\mu\left(\cup_{k \geq M} W_k^\dagger\right) \geq \epsilon\right) = I + II.$$

We now bound I . Observe that by Lemma 4.2,

$$\mu\left(W_l^\dagger\right) \langle \cosh(y) \rangle' \geq \left(1 - \frac{C}{\sqrt{N}}\right) \mu'(\tilde{W}_l)$$

with high probability. Thus I is bounded by

$$\begin{aligned}
I &\leq P(\mu'(\tilde{W}_l) \leq 2\epsilon \cdot L) + P(\langle \cosh(y) \rangle' \geq L) + o_N(1) \\
&\leq P(\mu'(\tilde{W}_l) \leq 2\epsilon \cdot L) + \frac{C(\xi')}{L} + o_N(1),
\end{aligned}$$

for each $L \geq 1$, where we have applied the localization lemma (Lemma A.5) in the second inequality.

We now turn to II . Observe that again by Lemma 4.2, with high probability,

$$\mu\left(\cup_{k \geq M} W_k^\dagger\right) \leq \left(1 + \frac{C}{\sqrt{N}}\right) \tilde{K}(\mu') \sqrt{\mu'(\cup_{k \geq M} \tilde{W}_k)}.$$

Thus for N sufficiently large,

$$\begin{aligned}
II &\leq P\left(2\tilde{K}(\mu') \sqrt{\mu'(\cup_{k \geq M} \tilde{W}_k)} \geq \epsilon\right) + o_N(1) \\
&\leq P(\mu'(\cup_{k \geq M} \tilde{W}_k) \geq \frac{\epsilon^2}{4L^2}) + P\left(\tilde{K}(\mu') \geq L\right) + o_N(1) \\
&= P(\mu'(\cup_{k \geq M} \tilde{W}_k) \geq \frac{\epsilon^2}{4L^2}) + \frac{C(\xi')}{L} + o_N(1)
\end{aligned}$$

where again in the last step we used Lemma A.5. Writing

$$\mu'(\tilde{W}_k) = v_k^N,$$

we can write the above as

$$I + II \leq P(v_l^N \leq 2\epsilon \cdot L) + P\left(\sum_{k \leq M} v_k^N \leq 1 - \frac{\epsilon^2}{4L^2}\right) + \frac{C}{L} + o_N(1).$$

Observe that the sets in the first two terms are closed in \mathcal{P}_m . Thus by the Portmanteau theorem and the fact that $(v_l^N) \rightarrow (v_l)$ in law on \mathcal{P}_m where (v_l) are $PD(\theta)$ with $\theta = 1 - \zeta(\{q_*\})$, we have that

$$\overline{\lim}_N I + II \leq P(v_l \leq 2\epsilon \cdot L) + P\left(\sum_{k \geq M} v_k \geq \frac{\epsilon^2}{4L^2}\right) + \frac{C}{L}.$$

We used here that for the Poisson-Dirichlet distribution $\sum v_k = 1$.

The Poisson-Dirichlet distribution satisfies

$$\mathbb{E} \sum_{k \geq M} v_k \leq f(M, \theta)$$

where $f \rightarrow 0$ as $M \rightarrow \infty$. In particular, by Markov's inequality we have

$$P\left(\sum_{k \geq M} v_k \geq \frac{\epsilon^2}{4L^2}\right) \leq \frac{4L^2}{\epsilon^2} f(M, \theta).$$

Thus combining the above we have that

$$\overline{\lim}_{M \rightarrow \infty} \overline{\lim}_{N \rightarrow \infty} P(\pi_N^{-1}(n) \geq M) \leq nP(v_n \leq 2\epsilon \cdot L) + n\frac{C}{L},$$

where we have used here that $v_n < v_k$ for $k < n$. Sending $\epsilon \rightarrow 0$ and then $L \rightarrow \infty$ and using the fact that $P(v_n = 0) = 0$, yields the result. \square

6.2 Proof of Theorem 1.1

Recall the notation

$$\langle \cdot \rangle_{\alpha, N}^T = (\mu')_N^T(\cdot | \tilde{W}_\alpha)$$

from Section 3 and recall that $\langle \cdot \rangle_{\alpha, N} = \mu_N(\cdot | C_\alpha)$. We begin by stating the following two lemmas whose proofs we will defer to the end of the section.

Lemma 6.4. *For every $\alpha \in \mathbb{N}$,*

$$\left| \langle s \rangle_{\alpha, N} - \langle s \rangle_{\pi_N^{-1}(\alpha), N}^T \right| \rightarrow 0$$

in probability as $N \rightarrow \infty$.

Lemma 6.5. *For every $\alpha \in \mathbb{N}$,*

$$\left| \tanh\left(\langle y \rangle_{\alpha, N} - (\xi'(1) - \xi'(q_*)) \langle s \rangle_{\alpha, N}\right) - \tanh\left(\langle y \rangle_{\pi_N^{-1}(\alpha), N}^T - (\xi'(1) - \xi'(q_*)) \langle s \rangle_{\pi_N^{-1}(\alpha), N}^T\right) \right| \rightarrow 0 \quad (6.2)$$

in probability as $N \rightarrow \infty$.

Proof of Theorem 1.1. By the above two lemmas, it suffices to prove (1.7) with $\langle \cdot \rangle_{\pi_N^{-1}(\alpha), N}^T$ replacing $\langle \cdot \rangle_{\alpha, N}$.

Now, let $Y_\alpha^N = \langle s \rangle_{\pi_N^{-1}(\alpha), N}^\top - \tanh \left(\langle y \rangle_{\pi_N^{-1}(\alpha), N}^\top - (\xi'(1) - \xi'(q_*)) \langle s \rangle_{\pi_N^{-1}(\alpha), N}^\top \right)$. Note that Y_α^N can be written as $Y_\alpha^N = X_{\pi(\alpha)}^N$ where

$$X_\alpha^N := \langle s \rangle_{\alpha, N}^\top - \tanh \left(\langle y \rangle_{\alpha, N}^\top - (\xi'(1) - \xi'(q_*)) \langle s \rangle_{\alpha, N}^\top \right),$$

By Lemma 6.1 and Lemma 6.3, it thus suffices to prove convergence of X_α^N to zero.

Observe that for X_α , this is a statement about a cavity coordinate with the local field independent of the measure μ' . Indeed, the Hamiltonian H' satisfies (3.1), and y satisfies (3.2). Thus, X_α^N goes to zero in probability by Corollary 3.2 and Theorem 1.1 follows. \square

We now turn to the proofs of the lemmas. Set

$$\langle \cdot \rangle_{\alpha, N} = \mu_N(\cdot | W_\alpha).$$

Proof of Lemma 6.4. We begin by observing that

$$\left| \langle s \rangle_{\alpha, N} - \langle \tilde{s} \rangle_{\alpha, N} \right| = \left| \int_{C_\alpha} s d\mu_N - \int_{W_\alpha} s d\mu_N \right| \leq \frac{2\mu_N(W_\alpha \Delta C_\alpha)}{\mu_N(C_\alpha)}$$

on the event that W_α and C_α both have positive mass. Since $\mu_N(W_\alpha \Delta C_\alpha) \rightarrow 0$ in probability by the essentially uniqueness theorem (Corollary 5.2) and $\mu_N(C_\alpha)$ and $\mu_N(W_\alpha)$ converge in law to a random variable that is almost surely positive, this goes to zero in probability. Then note that by the tilting lemma,

$$\left| \langle \tilde{s} \rangle_{\alpha, N} - \langle s \rangle_{\pi_N^{-1}(\alpha)}^\top \right| \leq \frac{C'}{\sqrt{N}}$$

with high probability, so that this too goes to zero in probability. The result then follows by the triangle inequality. \square

Proof of Lemma 6.5. As $\tanh(x)$ is 1-Lipschitz, and we know from Lemma 6.4 that $\langle s \rangle_\alpha - \langle \tilde{s} \rangle_{\pi_N^{-1}(\alpha)}^\top \rightarrow 0$ in probability, it suffices to show that

$$\langle y \rangle_{\alpha, N} - \langle y \rangle_{\pi_N^{-1}(\alpha), N}^\top \rightarrow 0$$

in probability. Observe that

$$\left| \int_{C_\alpha} y d\mu_N - \int_{W_\alpha} y d\mu_N \right| \leq \frac{1}{\mu_N(C_\alpha)} \|y\|_{L^2(\mu)} \sqrt{\mu_N(W_\alpha \Delta C_\alpha)} \left(1 + \frac{1}{\sqrt{\mu_N(W_\alpha)}} \right),$$

and that with probability tending to 1, $(1 + \frac{1}{\sqrt{\mu_N(W_\alpha)}}) \vee \frac{1}{\mu_N(C_\alpha)}$ will be finite. Furthermore, $\mu_N(W_\alpha \Delta C_\alpha) \rightarrow 0$ in probability by the quasi-uniqueness theorem (Theorem 5.1), and $\mathbb{E}\|y\|_2 \leq C$ uniformly in N by item 3 of Lemma A.5. Thus this tends to zero in probability as before. Similarly

$$\left| \langle \tilde{y} \rangle_{\alpha, N} - \langle y \rangle_{\pi_N^{-1}(\alpha)}^\top \right| \leq \frac{C'}{\sqrt{N}} \|y\|_{L^2(\mu')}$$

which goes to zero in probability by the same argument. \square

7 Evaluation of spin statistics

In this section, which is independent of the rest of the paper, we discuss how to calculate spin statistics. Using spin distributions, one can obtain formulae for expectations of products of spins, either through the directing function σ or by taking limits of expressions using Ruelle Cascades. The goal of this section is to explain how one can obtain expressions for such statistics as the solutions of certain partial differential equations. The input required will be the overlap distribution $\zeta(t)$. In particular, one can in principle evaluate these expression using standard methods from PDEs or numerically. Rather than developing a complete calculus of spin statistics, we aim to give a few illustrative examples.

At the heart of these calculations are the following key observation: the magnetization process for any finite collection $(\sigma^i)_{i=1}^n$ is a family of branching martingales whose independence properties mimics that of the tree encoding of their overlap arrays. (This can be formalized using the language of Branchingales. See [5] for more on this.) In this section we focus on two examples: two spin statistics, i.e., the overlap, and three spin statistics. One can of course write out general formulas, however, we believe that these two cases highlight the key ideas. In particular, the second case is the main example in [14], where this is calculated using replica theory. The reader is encouraged to compare the PDE and Martingale based discussion here with the notion of “tree operators” in that paper. For the remainder of this subsection, all state measures should be taken with boundary data $f(\sigma) = X_{q_*}(\sigma)$.

7.1 Two Spin Statistics

We first aim to study two spin statistics. As the spins take values ± 1 , there is only one nontrivial two spin statistic, namely $\mathbb{E}s_1^1 s_1^2$ where the subscript denotes the site index and the superscript denotes the replica index. Observe that by (1.16), we have that

$$\begin{aligned}\mathbb{E}s_1^1 s_1^2 &= \mathbb{E} \int \langle s \rangle_{\sigma^1} \cdot \langle s \rangle_{\sigma^2} d\mu^{\otimes 2} \\ &= \mathbb{E} \int \mathbb{E} \prod_{i=1}^2 u_x(q_*, X_{q_*}^{\sigma^i}) d\mu^{\otimes 2}.\end{aligned}$$

Observe that it suffices to compute $\mathbb{E}u_x(q_*, X_{q_*}^{\sigma^1})u_x(q_*, X_{q_*}^{\sigma^2})$. There are a few natural ways to compute this. Let $q_{12} = (\sigma^1, \sigma^2)$. One method is to observe that if $\Phi = \Phi_{q_{12}}$ solves

$$\begin{cases} (\partial_t + L_t^{lf})\Phi = 0 & [0, 1] \times \mathbb{R}^2 \\ \Phi(1, x, y) = \tanh(x) \tanh(y) \end{cases},$$

where L^{lf} is the infinitesimal generator for the local field process (see (A.1)), then

$$\mathbb{E}u_x(q_*, X_{q_*}^{\sigma^1})u_x(q_*, X_{q_*}^{\sigma^2}) = \Phi_{q_{12}}(0, h).$$

One can study this problem using PDE methods or Ito’s lemma. This yields the expression

$$\mathbb{E}s_1^1 s_1^2 = \int \Phi_s(0, h) d\zeta(s).$$

Alternatively, note that, by the branching martingale property of the magnetization process, we have that

$$\mathbb{E}u_x(q_*, X_{q_*}^{\sigma^1})u_x(q_*, X_{q_*}^{\sigma^2}) = \mathbb{E}u_x(q_{12}, X_{q_{12}})^2,$$

yielding the alternative expression

$$\mathbb{E}s_1^1 s_1^2 = \int \mathbb{E}u_x^2(s, X_s) d\zeta(s).$$

In the case that ζ is the Parisi measure for ξ (for this notation see [3, 11]), it is well-known that on the support of ζ ,

$$\mathbb{E}u_x^2(s, X_s) = s$$

so that

$$\mathbb{E}s_1^1 s_1^2 = \int s d\zeta(s).$$

This resolves a question from [7, Remark 5.5].

7.2 Three spin statistics.

We now turn to computing more complicated statistics. We focus on the case of the three spin statistic,

$$\mathbb{E}s_1^1 s_1^2 s_1^3,$$

as we believe this to be illustrative of the essential ideas.

We say a function $f : [0, 1]^k \rightarrow \mathbb{R}$ is *symmetric* if for every $\pi \in S_k$,

$$f(x_{\pi(1)}, \dots, x_{\pi(k)}) = f(x_1, \dots, x_k)$$

In the following, we denote by $dQ(R^n)$ the law of the overlap array $R^n = (R_{ij})_{ij \in [n]}$. We say that such a function has *vanishing diagonal* if $f(x, \dots, x) = 0$. We will always assume that Q satisfies the Ghirlanda-Guerra identities.

Our goal is to prove the following:

Theorem 7.1. *We have that*

$$\mathbb{E}s_1^1 s_1^2 s_1^3 = \frac{3}{4} \int \int \mathbb{E}u_x(b \vee a, X_{b \vee a})^2 u_x(a \wedge a, X_{a \wedge a}) d\zeta(a) d\zeta(b).$$

As a starting point, again observe that from the properties of state measures, (1.15),

$$\mathbb{E}s_1^1 s_1^2 s_1^3 = \mathbb{E} \int \langle s \rangle_{\sigma^1} \cdot \langle s \rangle_{\sigma^2} \cdot \langle s \rangle_{\sigma^3} d\mu^{\otimes 3}.$$

Denote the integrand by

$$\mathcal{R}(\sigma^1, \sigma^2, \sigma^3) = \langle s \rangle_{\sigma^1} \cdot \langle s \rangle_{\sigma^2} \cdot \langle s \rangle_{\sigma^3}.$$

The proof of this result will follow from the following two lemmas.

Lemma 7.2. *We have the following.*

1. Suppose that $g(x, y)$ is a continuous, symmetric function. Then

$$\int g(R_{12}, R_{13})dQ = \frac{1}{2} \int \int g(x, y)d\zeta(x)d\zeta(y) + \int g(x, x)d\zeta(x).$$

2. Suppose that $f(x, y, z)$ is a continuous symmetric function with vanishing diagonal. Then

$$\int f(R_{12}, R_{13}, R_{23})dQ = \frac{3}{2} \int f(R_{12} \vee R_{13}, R_{12} \wedge R_{13}, R_{12} \wedge R_{13})dQ.$$

3. Suppose that $f(x, y, z)$ is as above and such that $h(x, y) = f(x \vee y, x \wedge y, x \wedge y)$ is continuous. Then

$$\int f(R_{12}, R_{13}, R_{23})dQ = \frac{3}{4} \int \int h(x, y)d\zeta(x)d\zeta(y).$$

Proof. The last item is implied by the first two. We start with the second claim. Notice that there are three ways to choose a maximum of R_{12} , R_{13} , and R_{23} , so by symmetry of f and ultrametricity we have that

$$\int f(R_{12}, R_{13}, R_{23})dQ = 3 \int_{R_{12} > R_{13}} f(R_{12}, R_{13}, R_{13})dQ + \int_{R_{12} = R_{13} = R_{23}} f(R_{12}, R_{12}, R_{12})dQ$$

The second term is zero by the vanishing diagonal property of f , so that, since the integrand and event in the first term now no longer depends on the third coordinate, we have

$$RHS = 3 \int_{R_{12} \geq R_{13}} f(R_{12}, R_{13}, R_{13})dQ = \frac{3}{2} \int h(R_{12}, R_{13})dQ,$$

using again vanishing diagonal and the definition of h .

It remains to check the first claim. To this end, recall the Ghirlanda-Guerra identities

$$\int f(R_{12})g(R_{13})dQ = \frac{1}{2} \left[\int \int f(x)g(y)d\zeta(x)d\zeta(y) + \int f(x)g(x)d\zeta(x) \right].$$

By density of products, it follows for all $\phi = \phi(x, y)$ continuous that

$$\int \phi(x, y)dQ = \frac{1}{2} \left[\int \int \phi(x, y)d\zeta(x)d\zeta(y) + \int \phi(x, x)d\zeta(x) \right].$$

□

Lemma 7.3. *There is a continuous, symmetric function of three variables defined on the set of ultrametric $[0, 1]^3$ such that the function $\mathcal{R}(\sigma^1, \sigma^2, \sigma^3) = f(R_{12}, R_{13}, R_{23})$. This function has vanishing diagonal, and satisfies*

$$f(a, b, b) = \mathbb{E}u_x(b, X_b)^2 u_x(a, X_a) \tag{7.1}$$

for $a \leq b$.

Remark 7.4. This is to be compared with [14, Eq. 34].

Proof. That it is a continuous, symmetric function of the overlaps is obvious. It suffices to show (7.1). To this end, observe that without loss of generality $R_{12} \geq R_{13} = R_{23}$. In this case, denoting $R_{12} = b$ and $R_{23} = R_{13} = a$, we have that

$$\begin{aligned}\mathcal{R}(\sigma^1, \sigma^2, \sigma^3) &= \mathbb{E}u_x(1, X^{\sigma^1})u_x(1, X^{\sigma^2})u_x(1, X^{\sigma^3}) \\ &= \mathbb{E}u_x(b, X_b^1)u_x(b, X_b^2)u_x(b, X_b^3) \\ &= \mathbb{E}u_x(b, X_b^1)^2u_x(b, X_b^3) \\ &= \mathbb{E}u_x(b, X_b^1)^2u_x(a, X_a^3) \\ &= \mathbb{E}u_x(b, X_b)^2u_x(a, X_a).\end{aligned}$$

In the second line, we used independence and the martingale property. In the third line we used that the driving motions are identical in distribution until that time. In the fourth line we use the martingale property and independence of local fields again. The final result comes from the fact that the driving process for the three spins is equivalent until a . \square

We can now prove the main result of this subsection:

Proof of Theorem 7.1 . Recall that

$$\mathbb{E}s_1^1s_1^2s_3^3 = \mathbb{E}\langle \mathbb{E}\mathcal{R}(\sigma^1, \sigma^2, \sigma^3) \rangle.$$

The result then follow by combining Lemma 7.3 and part 3. of Lemma 7.2. \square

A Appendix

A.1 The clusters $C_{\alpha, N}$ and approximate ultrametricity

In this short section we summarize the properties of the clusters $C_{\alpha, N}$ used to construct the measures $\langle \cdot \rangle_{\alpha, N}$. These properties are described in the following theorem, which is a rephrasing of the main results in [10], specifically as in Section 9, Proposition 9.5-6 and Corollary 9.7 of that paper.

Theorem A.1 (Jagannath [10]). *Assume that $\zeta(\{q_*\}) > 0$. Then there are sequences $q_N \uparrow q_*$, ϵ_N, a_N, b_N all converging monotonically to 0, such $m_N \rightarrow \infty$, $m_N \cdot b_N^{1/8} \rightarrow 0$, $q_N + a_N < q_*$ and*

$$\zeta_N[q_N + a_N, 1] \geq \zeta(\{q_*\}) - b_N$$

for N sufficiently large and such that with probability $1 - o_N(1)$, there exist disjoint random sets $\{C_{\alpha, N}\}_{\alpha \in \mathbb{N}}$ of Σ_N :

1. *The collection $C_{\alpha, N}$ exhaust the set Σ_N :*

$$\sum_{\alpha} \mu_N(C_{\alpha, N}) \geq 1 - \epsilon_N.$$

2. *For any α , points are uniformly close:*

$$\mu_N^{\otimes 2}(\sigma^1, \sigma^2 \in C_{\alpha, N} : R(\sigma^1, \sigma^2) \leq q_N - a_N) \leq b_N.$$

3. For any $\alpha \neq \beta$,

$$\mu_N^{\otimes 2}(\sigma^1 \in C_{\alpha,N}, \sigma^2 \in C_{\beta,N} : R(\sigma^1, \sigma^2) \geq q_N + a_N) \leq b_N.$$

4. Uniformly in α we have,

$$\int_{C_{\alpha,N}^{\otimes 2}} |R_{12} - q_*| d\mu_N^2 < o_N(1).$$

5. The weights $\mu_N(C_{\alpha,N})$ converge to the weights of a Poisson-Dirichlet process of parameter $1 - \zeta(\{q_*\})$.

Note: We may always take $\alpha_{N-1} \geq N^{-1}$ in the above by monotonicity. That we can, take $m_N \cdot b_N^{1/8} \rightarrow 0$, follows by adding a constant to the definition of n_0 in Lemma 5.2 of [10].

A.2 On the driving process and its descendants

We record here the following basic properties of the driving process, cavity field process, local field process, and magnetization process.

Lemma A.2. *Let U be a positive ultrametric subset of a separable Hilbert space that is weakly closed and norm bounded equipped with the restriction of the Borel sigma algebra. Let $B_t(\sigma)$ be the process defined in (1.11). We have the following:*

1. *The covariance structure is positive semi-definite.*
2. *There is a version of this process that is jointly measurable and continuous in time.*
3. *For each σ , $B_t(\sigma)$ has the law of a brownian motion so that stochastic integration with respect to $B_t(\sigma)$ is well-defined.*

Proof. We begin with the first. To see this, simply observe that if $\alpha_i \in \mathbb{R}$, (t_i, σ_i) are finitely many points in $[0, q_*] \times U$ and $\sigma_* \in U$, then

$$\begin{aligned} \sum \alpha_i \alpha_j (t_i \wedge t_j \wedge (\sigma_i, \sigma_j)) &= \sum \alpha_i \alpha_j \int \mathbb{1}\{s \leq t_i\} \mathbb{1}\{s \leq t_j\} \mathbb{1}\{s \leq (\sigma_i, \sigma_j)\} ds \\ &\geq \sum \alpha_i \alpha_j \int \mathbb{1}\{s \leq t_i\} \mathbb{1}\{s \leq t_j\} \mathbb{1}\{s \leq (\sigma_i, \sigma_*)\} \mathbb{1}\{s \leq (\sigma_j, \sigma_*)\} ds \\ &= \|\sum \alpha_i \mathbb{1}\{s \leq t_i \wedge (\sigma_i, \sigma_*)\}\|_{L^2}^2 \geq 0. \end{aligned}$$

We now turn to the second. Observe first that, since $[0, q_*] \times U$ is separable and \mathbb{R} is locally compact, $B_t(\sigma)$ has a separable version. Furthermore, observe that $B_t(\sigma)$ is stochastically continuous in norm, that is as $(t, \sigma) \rightarrow (t_0, \sigma_0)$ in the norm topology, $P(|B_t(\sigma) - B_{t_0}(\sigma_0)| > \epsilon) \rightarrow 0$. Thus since U is weakly-closed and norm bounded it is compact in the weak topology. Thus it has a version that is jointly measurable by [9, Theorem IV.4.1]. Note then, since the covariance of $B_t(\sigma)$ for fixed σ is that of Brownian motion and $B_t(\sigma)$ is separable, it is in fact continuous by [9, Theorem IV.5.2].

The third property was implicit in the proof of the second. \square

We now observe the following consequence of the above proposition:

Corollary A.3. *Let U be a positive ultrametric subset of a separable Hilbert space that is weakly closed and norm bounded. Then the cavity field process, $Y_t(\sigma)$, the local field process, $X_t(\sigma)$, and the local magnetization process, $M_t(\sigma)$, exist, are continuous in time and Borel measurable in σ .*

In the above, the following observation regarding the infinitesimal generator of the above processes will be of interest.

Lemma A.4. *Let $(\sigma^i)_{i=1}^n \subset U$ where U is as above. Then we have the following.*

1. *The driving process satisfies the bracket relation*

$$\langle B(\sigma^1), B(\sigma^2) \rangle_t = \begin{cases} t & t \leq (\sigma, \sigma') \\ 0 & t > (\sigma, \sigma') \end{cases}.$$

2. *The cavity field process satisfies the bracket relation*

$$\langle Y(\sigma^1), Y(\sigma^2) \rangle_t = \begin{cases} \xi'(t) & t \leq (\sigma^1, \sigma^2) \\ 0 & \text{else} \end{cases}.$$

3. *The local fields process satisfies the bracket relation*

$$\langle X(\sigma^1), X(\sigma^2) \rangle_t = \begin{cases} \xi'(t) & t \leq (\sigma^1, \sigma^2) \\ 0 & \text{else} \end{cases}$$

and has infinitesimal generator

$$L_t^{lf} = \frac{\xi''(t)}{2} \left(\sum a_{ij}(t) \partial_i \partial_j + 2 \sum b_i(t, x) \partial_i \right) \quad (\text{A.1})$$

where $a_{ij}(t) = \mathbb{1} \{t \leq (\sigma^i, \sigma^j)\}$ and $b_i(t, x) = \zeta([0, t]) \cdot u_x(t, x)$.

Proof. We begin with the first claim. To see this, observe that by construction,

$$B_t(\sigma^1) = B_t(\sigma^2)$$

for $t \leq (\sigma^1, \sigma^2)$, thus the bracket above is just the bracket for Brownian motion. If $t > (\sigma^1, \sigma^2) := q$, then the increments $B_t(\sigma^1) - B_q(\sigma^1)$ and $B_t(\sigma^2) - B_q(\sigma^2)$ are independent Brownian motions. This yields the second regime. By elementary properties of Itô processes, we obtain the brackets for Y_t and X_t from this argument. It remains to obtain the infinitesimal generator for the local fields process.

To this end, observe that if $f = f(t, x_1, \dots, x_k)$ is a test function, then Itô's lemma applied to the process $(X_t(\sigma^i))_{i=1}^n$ yields

$$\begin{aligned} df &= \partial_t f \cdot dt + \sum_i \partial_{x_i} f \cdot dX_t(\sigma^i) + \frac{1}{2} \cdot \sum \partial_{x_i} \partial_{x_j} f \cdot d \langle X_t(\sigma^i), X_t(\sigma^j) \rangle \\ &= \left(\partial_t f + \sum_i \partial_{x_i} f \cdot (\xi''(t) \zeta(t) u_x(t, X_t(\sigma^i)) + \frac{\xi''}{2} \sum \mathbb{1} \{t \leq (\sigma^i, \sigma^j)\} \partial_{x_i} \partial_{x_j} f \right) dt + dMart \end{aligned}$$

where $dMart$ is the increment for some martingale. Taking expectations and limits in the usual fashion then yields the result. \square

A.3 Tail bounds for some Gibbs averages

Lemma A.5 (Localization Lemma). *Recall $\tilde{K}(\mu')$ from (4.4), $y_{\alpha,N}$ from Section 3 and y_N from (1.5). For any $L > 0$ we have the following estimates.*

1. For any $\alpha \in \mathbb{N}$,

$$P(|y_{\alpha,N}| > L) \leq C_1(\xi') \cdot e^{-C_2(\xi')L^2}.$$

2. For any $\alpha \in \mathbb{N}$,

$$P\left(\int_{X_\alpha} \cosh(y_{\alpha,N}) d\nu_N > L\right) \leq C(\xi')/L.$$

3. We have that

$$P(\tilde{K}(\mu') \geq L) \leq C(\xi')/L.$$

4. We have that

$$\mathbb{E}\left(\int_{\Sigma_N} y_N^2 d\mu_N\right)^{1/2} \leq C(\xi').$$

Proof. In the following we will drop the index α of our notation without any loss. To see the first item, note that $y_{\alpha,N}$ has finite moment generating function. Fix $\lambda \geq 1$. We have

$$\begin{aligned} \mathbb{E}e^{\lambda y_N} &= \mathbb{E}\left[\frac{\int_{X_\alpha} e^{\lambda y_N(\sigma)} \cosh(y_N(\sigma)) d\nu_N}{\int_{X_\alpha} 2 \cosh(y_N(\sigma)) d\nu_N}\right] + P(X_\alpha = \emptyset) \\ &\leq \mathbb{E}\left[\int_{X_\alpha} e^{\lambda y_N(\sigma)} \cosh(y_N(\sigma)) d\nu_N\right] + 1 \\ &\leq \mathbb{E} \int_{X_\alpha} \mathbb{E}(\exp(\lambda y_N) \cosh(y_N) | \nu_N) d\nu_N + 1 \\ &= \frac{1}{2} \left(e^{(1+\lambda)^2 \xi'(1)}\right) + 1. \end{aligned} \tag{A.2}$$

Then, by Markov's inequality, we have

$$P(y_N \geq L) \leq \mathbb{E}e^{\lambda y_N - \lambda L} \leq \mathbb{E}e^{(1+\lambda)^2 \xi'(1) - \lambda L} \leq C_1(\xi') \cdot e^{-C_2(\xi')L^2},$$

for L sufficiently large by choosing $\lambda = L/2$, for instance. Increasing the value of $C_1(\xi')$ if necessary we obtain the result for all $L > 0$. Similarly for $-y_N$.

The second item holds by Markov's inequality, conditioning on ν_N and using the Gaussian bound of item 1. For the third item, note that using Lemma A.6, conditioning on μ' and letting Z be a Gaussian random variable with variance $\xi'(1)$, we have

$$P(\tilde{K}(\mu') \geq L) \leq C(\xi') \mathbb{E}[e^{-4Z} \cosh Z] \mu'_N(\Sigma_N) \leq L^{-1} C(\xi'),$$

as desired.

We prove the last item as follows. To see this observe that it suffices to bound $\mathbb{E} \int y^2 d\mu$. To estimate this, observe that if $\Delta = \max |r(1, \sigma) - r(-1, \sigma)|$, then

$$\begin{aligned} \mathbb{E} \int y^2 d\mu &\leq \mathbb{E} e^{2\Delta} \frac{\int_{\Sigma_{N-1}} y^2 \cosh(y) d\mu'}{\int_{\Sigma_{N-1}} \cosh(y) d\mu'} \\ &\leq (\mathbb{E} e^{4\Delta})^{1/2} \left(\mathbb{E} \int y^4 \cosh(y) d\mu' \right)^{1/2}, \end{aligned}$$

where in the last inequality we use Cauchy-Schwarz and the fact that $\cosh(x) \geq 1$. Observe that the first term is bounded by (A.4). Since y is independent of μ' , we can integrate in y to find that the second term is also uniformly bounded. \square

A.4 Decomposition and regularity of mixed p -spin Hamiltonians

In this section, we present some basic properties of Mixed p -spin Hamiltonians. Recall that for $\sigma = (\sigma_1, \dots, \sigma_N) \in \Sigma_N$, $\rho(\sigma) = (\sigma_2, \dots, \sigma_N) \in \Sigma_{N-1}$. Now observe that for any mixed p -spin glass model, the Hamiltonian has the following decomposition:

$$H_N(\sigma) = \tilde{H}_N(\rho(\sigma)) + \sigma_1 y_N(\rho(\sigma)) + r_N(\sigma_1, \rho(\sigma)), \quad (\text{A.3})$$

where the processes come from the following lemma.

Lemma A.6. *There exist centered Gaussian processes \tilde{H}_N, y_N, r_N such that (4.1) holds and*

$$\begin{aligned} \mathbb{E} \tilde{H}_N(\sigma^1) \tilde{H}_N(\sigma^2) &= N \xi \left(\frac{N-1}{N} R_{12} \right), \\ \mathbb{E} y_N(\sigma^1) y_N(\sigma^2) &= \xi'(R_{12}) + o_N(1), \\ \mathbb{E} r_N(\sigma^1) r_N(\sigma^2) &= O(N^{-1}). \end{aligned}$$

Furthermore, there exist positive constant C_1 and C_2 so that with probability at least $1 - e^{-C_1 N}$,

$$\max_{\sigma \in \Sigma_{N-1}} |r_N(1, \sigma) - r_N(-1, \sigma)| \leq \frac{C_2}{\sqrt{N}},$$

and a positive constant C_3 so that

$$\mathbb{E} \exp \left(2 \max_{\sigma \in \Sigma_{N-1}} |r_N(1, \sigma) - r_N(-1, \sigma)| \right) \leq C_3. \quad (\text{A.4})$$

Proof. The lemma is a standard computation on Gaussian processes. To simplify the exposition we will consider the pure p -spin model. The general case follows by linearity. Here, we set

$$\tilde{H}_N(\rho(\sigma)) = N^{-\frac{p-1}{2}} \sum_{2 \leq i_1, \dots, i_p \leq N} g_{i_1 \dots i_p} \sigma_{i_1} \dots \sigma_{i_p},$$

$$y_N(\rho(\sigma)) = N^{-\frac{p-1}{2}} \sum_{k=1}^p \sum_{\substack{2 \leq i_1, \dots, i_p \leq N \\ i_k=1}} g_{i_1 \dots i_p} \sigma_{i_1} \dots \sigma_{i_p},$$

and

$$r_N(\sigma_1, \rho(\sigma)) = N^{-\frac{p-1}{2}} \sum_{l=2}^p \sigma_1^\ell \sum_{2 \leq i_1, \dots, i_{p-\ell} \leq N} J_{i_1 \dots i_{p-\ell}} \sigma_{i_1} \dots \sigma_{i_{p-\ell}},$$

where $J_{i_1 \dots i_{p-\ell}}$ are centered Gaussian random variables with variance equal to $\binom{p}{\ell}$: $J_{i_1 \dots i_{p-\ell}}$ is the sum of the $g_{i_1 \dots i_p}$ where the index 1 appears exactly ℓ times. Computing the variance of these three Gaussian processes give us the first three statements of the Lemma. For the second to last and last statement, note that for any $\sigma \in \Sigma_{N-1}$, $r(1, \sigma) - r(-1, \sigma)$ is a centered Gaussian process with variance equal to

$$\frac{4}{N^{p-1}} \sum_{\ell=3, \ell \text{ odd}}^p \binom{p}{\ell} (N-1)^{p-\ell} \leq \frac{C_p}{N^2},$$

for some constant C_p . A standard application of Borell's inequality (Theorem 2.1.1 in [1]), the tail estimate for the maximum of a Gaussian process (Equation (2.1.4) in [1]) and Sudakov-Fernique inequality (Theorem 2.2.3 in [1]) gives us the desired result. \square

A.5 The Cavity Equations and Ghirlanda-Guerra Identities

We say that ν in \mathcal{M} satisfies the Ghirlanda-Guerra identities if the law of the array (1.9) satisfies the following property: for every $f \in C([-1, 1]^n)$ and $g \in C([-1, 1])$, we have

$$\mathbb{E} \langle f(R^n) \cdot g(R_{1,n+1}) \rangle = \frac{1}{n} \left[\mathbb{E} \langle f(R^n) \rangle \cdot \mathbb{E} \langle g(R_{12}) \rangle + \sum_{k=2}^n \mathbb{E} \langle f(R^n) \cdot g(R_{1k}) \rangle \right], \quad (\text{A.5})$$

where by the bracket, $\langle \cdot \rangle$, we mean integration against the relevant products of μ with itself.

We say that ν satisfies the cavity equations if the following is true. Fix the directing function σ as in Section 1.2 and let $\bar{\sigma}$ be as above. Let $g_{\xi'}(\bar{\sigma})$ denote the centered Gaussian process indexed by $L^2([0, 1], dv)$ with covariance

$$\mathbb{E} g_{\xi'}(\bar{\sigma}(w, u, \cdot)) g_{\xi'}(\bar{\sigma}(w, u', \cdot)) = \xi' \left(\int \bar{\sigma}(w, u, v) \bar{\sigma}(w, u', v) dv \right)$$

and let $G'_\xi(\bar{\sigma}) = g_{\xi'}(\bar{\sigma}) + z(\xi'(1) - \xi'(\|\bar{\sigma}(w, u, \cdot)\|_{L^2(dv)}^2))^{1/2}$. Let $g_{\xi', i}$ and $G_{\xi', i}$ be independent copies of these processes. Let $n, m, q, r, l \geq 1$ be such that $n \leq m$ and $l \leq q$. Let $C_l \subset [m]$ and let $C_l^1 = C_l \cap [n]$ and $C_l^2 = C_l \cap (n + [m])$. Let

$$U_l = \int \mathbb{E}' \prod_{i \in C_l^1} \tanh G_{\xi', i}(\bar{\sigma}(w, u, \cdot)) \prod_{i \in C_l^2} \bar{\sigma}_i \mathcal{E}_{n, r} du$$

where \mathbb{E}' is expectation in z , $\bar{\sigma}_i = \bar{\sigma}(w, u, v_i)$, $\theta(t) = \xi'(t)t - \xi(t)$, and where

$$\mathcal{E}_{n, r} = \exp \left(\sum_{i \leq n} \log \cosh(G_{\xi', i}(\bar{\sigma}(w, u, \cdot))) + \sum_{k \leq r} G_{\theta, k}(\bar{\sigma}(w, u, \cdot)) \right).$$

Let $V = \mathbb{E}' \mathcal{E}_{n,r}$. The *cavity equations* are then given by

$$\mathbb{E} \prod_{l \leq q} \mathbb{E}' \prod_{i \in C_l} \bar{\sigma}_i = \mathbb{E} \frac{\prod_{l \leq q} U_l}{V^q} \quad (\text{A.6})$$

for all $n, m, q, r \geq 1$. We let \mathcal{M}_{inv}^ξ denote the space of ξ -cavity invariant measures that satisfy the Ghirlanda-Guerra identities.

A.6 Cavity equations for generic models

In this section, we discuss the cavity equations for Generic mixed p -spin models. The argument for the following theorem is fairly standard—see for example [17, Chapter 3, Theorem 3.6]—so we only provide a sketch here.

Proposition A.7. *The generic models satisfy the cavity equations (A.6) with $r = 0$.*

Proof Sketch. Fix n sites and a C_l as in (A.6). By site symmetry, we may assume that these are the last n sites. Our goal is then to show that

$$\mathbb{E} \prod_{l \leq q} \left\langle \prod_{i \in C_l} \sigma_i \right\rangle = \mathbb{E} \prod_{i \leq q} \frac{\left\langle \prod_{i \in C_l} \tanh(g_{\xi',i}(\sigma)) \mathcal{E}_n \right\rangle}{\langle \mathcal{E}_n \rangle} + o_N(1). \quad (\text{A.7})$$

With this observation in hand, note that if we write a configuration τ as $\tau = (\sigma, \epsilon)$ with $\sigma \in \Sigma_{N-n}$, $\epsilon \in \Sigma_n$, we have the analogue of (4.1) and Lemma A.6 with:

$$H_N(\tau) = \tilde{H}_N(\sigma) + \sum_{i=1}^n \epsilon_i y_{N,i}(\sigma) + r_N(\sigma, \epsilon).$$

This implies that the left side of (A.7) is equivalent to

$$\mathbb{E} \prod_{l \leq q} \frac{\left\langle \prod_{i \in C_l} \tanh(y_{N,i}(\sigma)) \mathcal{E}_{n,0} \right\rangle_{G'}}{\langle \mathcal{E}_{n,0} \rangle_{G'}^q},$$

where G' is the Gibbs measure for \tilde{H}_N on Σ_{N-n} .

By a localization and Stone-Weierstrass argument, we see that it suffices to show that

$$\mathbb{E} \prod_{l \leq q} \left\langle \prod_{i \in C_l} \tanh(y_i(\sigma)) \mathcal{E}_n \right\rangle_{G'} \langle \mathcal{E}_n \rangle_{G'}^k = \mathbb{E} \prod_{l \leq q} \left\langle \prod_{i \in C_l} \tanh(g_{\xi',i}(\sigma)) \mathcal{E}_n \right\rangle_G \langle \mathcal{E}_n \rangle_G^k + o_N(1).$$

Evidently, this will follow provided the limiting overlap distribution for $\mathbb{E} G'^{\otimes \infty}$ and $\mathbb{E} G^{\otimes \infty}$ are the same. As generic models are known to have a unique limiting overlap distribution (by Lemma 3.6 of [17]), it suffices to show that in fact the overlap distribution of law of H'_N and H_{N-n} are the same. Observe that

$$|\text{Cov}_{H'}(\sigma^1, \sigma^2) - \text{Cov}_H(\sigma^1, \sigma^2)| = N \left| \xi \left(\frac{N}{N+n} R_{12} \right) - \xi(R_{12}) \right| \lesssim_{\xi',n} 1,$$

uniformly for $\sigma^1, \sigma^2 \in \Sigma_{N-n}$, so that by a standard interpolation argument (see, e.g., [17, Theorem 3.6]) we have that the free energy of these two systems is the same in the limit $N \rightarrow \infty$. An explicit differentiation argument (see [17, Theorem 3.7]) combined with [17, Theorem 2.13] shows that the overlap distributions are the same. \square

References

- [1] Robert Adler and Jonathan Taylor. *Random fields and geometry*. Springer monographs in mathematics, 2007.
- [2] Louis-Pierre Arguin and Michael Aizenman. On the structure of quasi-stationary competing particle systems. *Ann. Probab.*, 37(3):1080–1113, 05 2009.
- [3] Antonio Auffinger and Wei-Kuo Chen. On properties of Parisi measures. *Probab. Theory Related Fields*, 161(3-4):817–850, 2015.
- [4] Antonio Auffinger and Wei-Kuo Chen. The Parisi formula has a unique minimizer. *Comm. Math. Phys.*, 335(3):1429–1444, 2015.
- [5] Tim Austin. Exchangeable random measures. *Ann. Inst. H. Poincaré Probab. Statist.*, 51(3):842–861, 08 2015.
- [6] Erwin Bolthausen. An iterative construction of solutions of the TAP equations for the Sherrington-Kirkpatrick model. *Comm. Math. Phys.*, 325(1):333–366, 2014.
- [7] Erwin Bolthausen and Alain-Sol Sznitman. On Ruelle’s probability cascades and an abstract cavity method. *Comm. Math. Phys.*, 197(2):247–276, 1998.
- [8] Sourav Chatterjee. Spin glasses and Stein’s method. *Probab. Theory Related Fields*, 148(3-4):567–600, 2010.
- [9] Iosif I. Gikhman and Anatoli V. Skorokhod. *The theory of stochastic processes. I*. Classics in Mathematics. Springer-Verlag, Berlin, 2004. Translated from the Russian by S. Kotz, Reprint of the 1974 edition.
- [10] Aukosh Jagannath. Approximate Ultrametricity for Random Measures and Applications to Spin Glasses. *Comm. Pure Appl. Math.*, To Appear.
- [11] Aukosh Jagannath and Ian Tobasco. Some Properties of the Phase Diagram for Mixed p -Spin Glasses. *Probab. Theory Related Fields*, To Appear.
- [12] Aukosh Jagannath and Ian Tobasco. A dynamic programming approach to the Parisi functional. *Proc. Amer. Math. Soc.*, 144(7):3135–3150, 2016.
- [13] Marc Mézard, Giorgio Parisi, and Miguel Angel Virasoro. *Spin glass theory and beyond*, volume 9. World scientific Singapore, 1987.
- [14] Marc Mézard and Miguel Angel Virasoro. The microstructure of ultrametricity. *Journal de Physique*, 46(8):1293–1307, 1985.
- [15] Dmitry Panchenko. The Ghirlanda-Guerra identities for mixed p -spin model. *C. R. Math. Acad. Sci. Paris*, 348(3-4):189–192, 2010.
- [16] Dmitry Panchenko. The Parisi ultrametricity conjecture. *Ann. of Math. (2)*, 177(1):383–393, 2013.
- [17] Dmitry Panchenko. *The Sherrington-Kirkpatrick model*. Springer, 2013.

- [18] Dmitry Panchenko. Spin glass models from the point of view of spin distributions. *The Annals of Probability*, 41(3A):1315–1361, 05 2013.
- [19] Dmitry Panchenko. Hierarchical exchangeability of pure states in mean field spin glass models. *Probab. Theory Related Fields*, 161(3-4):619–650, 2015.
- [20] Daniel W. Stroock and S. R. Srinivasa Varadhan. *Multidimensional diffusion processes*, volume 233. Springer Science & Business Media, 1979.
- [21] Michel Talagrand. *Spin glasses: a challenge for mathematicians*, volume 46. Springer-Verlag, Berlin, 2003. Cavity and mean field models.
- [22] Michel Talagrand. Construction of pure states in mean field models for spin glasses. *Probability Theory and Related Fields*, 148(3-4):601–643, 2010.
- [23] David J Thouless, Philip W Anderson, and Robert G Palmer. Solution of 'solvable model of a spin glass'. *Philosophical Magazine*, 35(3):593–601, 1977.